



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

**THE SHEATH FLOW CUVETTE
IN HIGH-SENSITIVITY FLUORESCENCE DETECTION
FOR DNA SEQUENCING
IN SINGLE AND MULTIPLE CAPILLARY SYSTEMS**

**BY
JIANZHONG ZHANG**



**A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of DOCTOR OF PHILOSOPHY**

DEPARTMENT OF CHEMISTRY

Edmonton, Alberta

SPRING 1994



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-11426-0

Canada

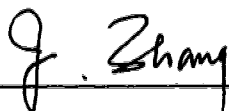
UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jianzhong Zhang
TITLE OF THESIS: The Sheath Flow Cuvette in High-Sensitivity
Fluorescence Detection for DNA Sequencing
in Single and Multiple Capillary Systems
DEGREE: Doctor of Philosophy
YEAR THIS DEGREE
GRANTED: 1994

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



10735-81 Avenue, Suite 201

Edmonton, Alberta T6E 1Y2

Date Jan. 28, 1994

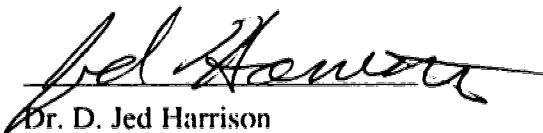
UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled THE SHEATH FLOW CUVETTE IN HIGH-SENSITIVITY FLUORESCENCE DETECTION FOR DNA SEQUENCING IN SINGLE AND MULTIPLE CAPILLARY SYSTEMS submitted by JIANZHONG ZHANG in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.



Dr. Norman J. Dovichi



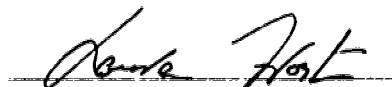
Dr. D. Jed Harrison



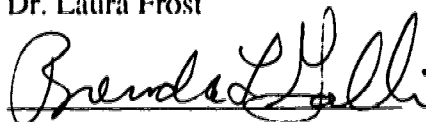
Dr. Liang Li



Dr. Ole Hindsgaul



Dr. Laura Frost



Dr. Brenda L. Gallie

January 21, 1994

Abstract

Capillary gel electrophoresis in DNA sequencing is shown to provide several advantages over conventional slab gel electrophoresis in terms of speed, resolution and efficiency. Laser-induced fluorescence detection incorporated with a sheath flow cuvette as post-column detector proved to be valuable for both single capillary and multiple capillary systems.

A single capillary system was built for DNA sequencing of samples in which four terminating dideoxynucleotides were labeled with a family of distinctive succinylfluorescein dyes. Fluorescence, induced by a single laser line at 488 nm, was detected in two spectral channels. The ratio of the fluorescence intensity in the two channels was used to identify the terminating nucleotide. Sequencing rates up to 1000 bases/h were obtained at an electric field strength of 465 V/cm in 4%T, 5%C polyacrylamide gel. Sequence was determined to 250 bases. The mass detection limit (3σ) of the system was less than 2 zmol fluorescein [1 zmol (zeptomole) = 10^{-21} mol = 600 molecules] in the two spectral channels.

A multiple capillary system increases the sequencing throughput and yet retains the performance of the single capillary system. At the first stage of its development, capillary zone electrophoresis of 15 amino acids and capillary gel electrophoresis of DNA samples were used to characterize the system. Although the migration rate of the same sample in different capillaries showed some difference, the signals from samples in different capillaries showed no cross-talk.

A five capillary system was configured as a two-laser-line, four-spectral-channel sequencer for Applied Biosystems, Inc. (ABI) samples produced with a set of four dye-labeled dideoxynucleotides. Non-crosslinked polyacrylamide polymer seems essential for DNA sequencing fragment separation in the multiple capillary format. Sequencing speed of over 2000 bases/h was achieved at 200 V/cm in non-cross linked polyacrylamide gel

by using this five capillary system. Readable sequences extended beyond 500 bases. The detection limits (3σ) were 130 ± 30 fluorescein molecules injected into the capillaries.

Table of Contents

Chapter 1

Introduction	1
1.1. DNA	1
1.2. DNA Sequencing	4
1.3. DNA Sequencing Approaches	8
1.3.1. One-Dye Primer (Pharmacia).....	8
1.3.2. Four Dye-Primers (ABI)	9
1.3.3. Four Dye-Terminators (DuPont And ABI)	9
1.3.4. Dye-Deoxy or Internal Label	13
1.3.5. One-Dye Peak-Height	13
1.3.6. Two-Dye, Peak-Height	14
1.3.7. Two-Dye Binary and One-Label Three-Heights	14
1.3.8. Cycle Sequencing.....	15
1.3.9. Primer Walking by Contiguous Hexamers	16
1.4. Capillary Zone Electrophoresis.....	19
1.5. Capillary Gel Electrophoresis	24
1.6. Laser-Induced Fluorescence Detection	25
References	31

Chapter 2

High-Speed, High-Sensitivity DNA Sequencing by Capillary Gel Electrophoresis Based on Sheath Flow Cuvette as A Fluorescence

Detector	34
2.1 Introduction	34
2.2. Instrumental	36
2.2.1. Design Consideration	36

2.2.2. Instrument	39
2.3. Experimental	44
2.3.1. Material	44
2.3.2. Capillary Gel Preparation.....	44
2.3.3. DNA Sequencing Sample	45
2.4. Results And Discussion	46
2.4.1. Limits of Detection (LOD)	46
2.4.2. DNA Sequence Determination.....	49
References	63

Chapter 3

Multiple Capillary System (1): A Preliminary Investigation	65
3.1. Introduction	65
3.2. Instrumental	67
3.2.1. Design Consideration	67
General	67
Capillary Quality Assurance	67
Power Supply	69
Sample Loading	69
Sample.....	69
Fluorescence Detection	69
3.2.2. Silicon Avalanche Photodiode (SiAPD) and Gradient Index (GRIN) Lens	71
SiAPD	71
GRIN Lens	72
3.2.3. Instrument	75
3.3. Experimental	80
3.3.1. Reagents	80

3.3.2. Methods.....	80
3.4. Results and Discussion.....	81
3.4.1. Detection Limits.....	81
3.4.2. Fluorescein Labeled Amino Acid Separation by Capillary Zone Electrophoresis.....	83
3.4.3. DNA Sequencing by Capillary Gel Electrophoresis.....	87
3.5. Conclusion	92
References	93

Chapter 4

Multiple Capillary System (2): High Performance DNA Sequencer	95
4.1. Introduction	95
4.2. Instrumental	96
4.2.1. Sheath Flow Profile.....	96
4.2.2. Laser Beam Profile.....	97
4.2.3. Fluorescence Detection Optical Train.....	102
4.3. Experimental	109
4.3.1. Material	109
4.3.2. Methods.....	112
4.4 Results and Discussion.....	112
4.4.1. Instrument Performance	112
4.4.2. Gel	119
4.4.3. DNA Sequencing	141
4. 5. Conclusions	154
References	157

Chapter 5

Conclusions and Future Work	160
5.1. Conclusions	160

5.2. Future Work	161
References	164

List of Figures

Fig. 1.1. (a) A schematic diagram of DNA structure.	2
Fig. 1.1. (b) Molecular structures of A, C, G and T.	3
Fig. 1.2. A schematic of sample generation by Sanger's method.	5
Fig. 1.3. The conventional DNA sequencing by gel electrophoresis.	7
Fig. 1.4. Structures of DuPont's dye-terminators.	11
Fig. 1.5. Four fluorescein dye-terminators used in ABI's PRISM system.	12
Fig. 1.6. Schematic diagrams of (a) PCR and (b) cycle sequencing.	17
Fig. 1.7. Schematic diagrams of (a) shotgun strategy and (b) primer walking strategy.	18
Fig. 1.8. Schematic diagram of a capillary electrophoresis system (post-column detector).	21
Fig. 1.9. The normalized absorbance spectrum of fluorescein.	28
Fig. 1.10. The normalized fluorescence spectrum of fluorescein (uncorrected).	29
Fig. 2.1. A schematic of the sheath flow cuvette as a post-column fluorescence detector.	38
Fig. 2.2. Sheath flow cuvette and two-spectral channel laser-induced fluorescence detection system.	40
Fig. 2.3. The spectra of the bandpass filter, 530DF45 (solid line), and the dichroic beamsplitter, 525DCLP (dashed line).	42
Fig. 2.4. Detection limit measurements in the two spectral channels.	47
Fig. 2.5. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection (287V/cm).	50
Fig. 2.6. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection (461 V/cm).	52

Fig. 2.7. Migration time as a function of fragment length for single-stranded DNA fragments separated in a 4%T, 5%T gel at different electric field strengths.	53
Fig. 2.8. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection (465 V/cm).	54
Fig. 2.9. A current increase as a function of recovery time.	56
Fig. 2.10. Current profile	56
Fig. 2.11. Expanded region corresponding to nucleotides 60-100 in Fig. 2.8	59
Fig. 2.12. The effect of polarized light on the spectrum of the dichroic beamplitter, 525DCLP	62
Fig. 3.1. A schematic of sheath flow cuvette based fluorescence detection chamber for an array of five capillaries	68
Fig. 3.2. A schematic of passive quenched circuit and the resulting pulses.	73
Fig. 3.3. Comparison of a GRIN rod lens with a conventional lens.	76
Fig. 3.4. The fluorescence detection system for the multiple capillaries.	78
Fig. 3.5. A counter block diagram.	79
Fig. 3.6. Calibration curves for the three capillary system.	82
Fig. 3.7. Simultaneous separation of 3 groups of amino acids	84
Fig. 3.8. The migration time variation of analyte in three capillaries.	86
Fig. 3.9. Three-capillary, one-spectral-channel sequencing of M13mp18. (a) extended run.	88
Fig. 3.9. Three-capillary, one-spectral-channel sequencing of M13mp18. (b) expanded region	91
Fig. 4.1. A tapered sheath flow cuvette for an array of five capillaries.	98
Fig. 4.2. A schematic of the transformation of a Gaussian beam by a thin lens.	100
Fig. 4.3. The laser beam profile.	103
Fig. 4.4. Dark counts and background counts of the photon counting modules	106
Fig. 4.5. The arrangement of GRIN lens array in the image plane	108

Fig. 4.6. (a) The image of the GRIN lens array	110
Fig. 4.6. (b) The image of superimposition of the fluorescence spots and back-illuminated spots.	111
Fig. 4.7. Injection of 1300 fluorescein molecules.	114
Fig. 4.8. Electropherogram of fluorescein	117
Fig. 4.9. The signal-to-noise ratio as a function of the laser power	118
Fig. 4.10. Electropherograms of FAM-labeled, A-terminated DNA sequencing fragments generated from M13mp18	121
Fig. 4.11. A schematic of capillary gel-filling apparatus.	125
Fig. 4.12. Electropherograms of FAM-labeled, C-terminated M13mp18 DNA sequencing fragments separated in five capillaries filled with 5%T linear polyacrylamide.	127
Fig. 4.13. The relationship between the eluting time of the reptation peaks and the average current of the electrophoresis	130
Fig. 4.14. Electropherograms of FAM-labeled, C-terminated M13mp18 DNA sequencing fragments separated in five capillaries filled with 4.0, 4.9, 5.7, 6.5, 7.3 %T linear polyacrylamide.	133
Fig. 4.15. Migration time as a function of DNA fragment length in different %T	136
Fig. 4.16. Mobility of FAM-labeled DNA fragments in different %T	137
Fig. 4.17. Ferguson plots for DNA fragments	139
Fig. 4.18. Mobility of the DNA fragment as a function of inverse fragment length in linear polyacrylamide with different T%.	140
Fig. 4.19. Laser-induced fluorescence detector for four spectral-channel sequencing.	143
Fig. 4.20. (a) The sequencing raw data and the selected maxima.....	144
Fig. 4.20. (b) A section of the processed data showing the four spectral information	145

Fig. 4.21. (a) Four spectral, dye-terminator DNA sequencing.	148
Fig. 4.21. (b) A section of simultaneous separation in five capillaries of DNA sequencing fragments.	155
Fig. 5.1. A schematic diagram of a sheath flow cuvette fluorescence detector for two-dimensional capillary array and four-spectral fluorescence detection	163

List of Tables

Table 2.1. Two-spectral-channel fluorescence detection of fluorescein molecules.....	46
Table 2.2. Some values KLOD of for different interval lengths.....	48
Table 3.1 Parameters relevant to PMT and SiAPD	71
Table 3.2. Data Sheet for SPCM-100-PQ (Selected)	77
Table 4.1. Collection Efficiencies of Different Lenses ($n = 1$).....	104

List of Abbreviations

ABI: Applied Biosystems, Inc.

Bis: N,N'-methylene bisacrylamide

CCD: charge coupled device

DNA: deoxyribonucleic acid

dNTP: deoxynucleoside triphosphate

ddNTP: dideoxynucleoside triphosphate

FITC: fluorescein isothiocyanate

GRIN: gradient index

He-Ne: Helium-neon

ID: inner diameter

LOD: limit of detection

NA: numerical aperture

OD: outer diameter

PCR: polymerase chain reaction

PMT: photomultiplier tube

SiAPD: silicon avalanche photodiode

SPCM: single photon counting module

TBE: tris-boric acid-EDTA

TEMED: N,N,N',N'-tetramethylethylenediamine

Tris: Tris(hydroxymethyl)aminomethane

μ: mobility

Chapter 1

Introduction

DNA sequencing by capillary gel electrophoresis coupled with fluorescence detection is an emerging technique that features small sample loading, high sequencing speed, and potential automation. This technique is also very challenging, for it requires high-sensitivity fluorescence detection, a very reliable separation medium and, for large-scale DNA sequencing projects, simultaneous handling of a large numbers of capillaries. Before describing the electrophoresis system, the basics of DNA sequencing are reviewed.

1.1. DNA

In 1953, Crick and Watson discovered the structure of that most important molecule—deoxyribonucleic acid or DNA [1] and led to the way to an understanding of life science [2]. According to their model, DNA is a long, linear molecule that is made of two strands; the two strands are antiparallel—with one strand in 3' → 5' direction and its complementary strand in 5' → 3' direction (Fig. 1.1). Along a DNA strand, four bases, adenine (A), cytosine (C), guanine (G), and thymine (T), are attached to a backbone of sugars and phosphates. The two strands are associated with each other through the hydrogen bonds: A on one strand is always associated with T on the other to form an A-T base pair; C on one strand is always associated with G on the other to form a C-G base pair. The sequence of bases along the sugar-phosphate backbone encodes the genetic information.

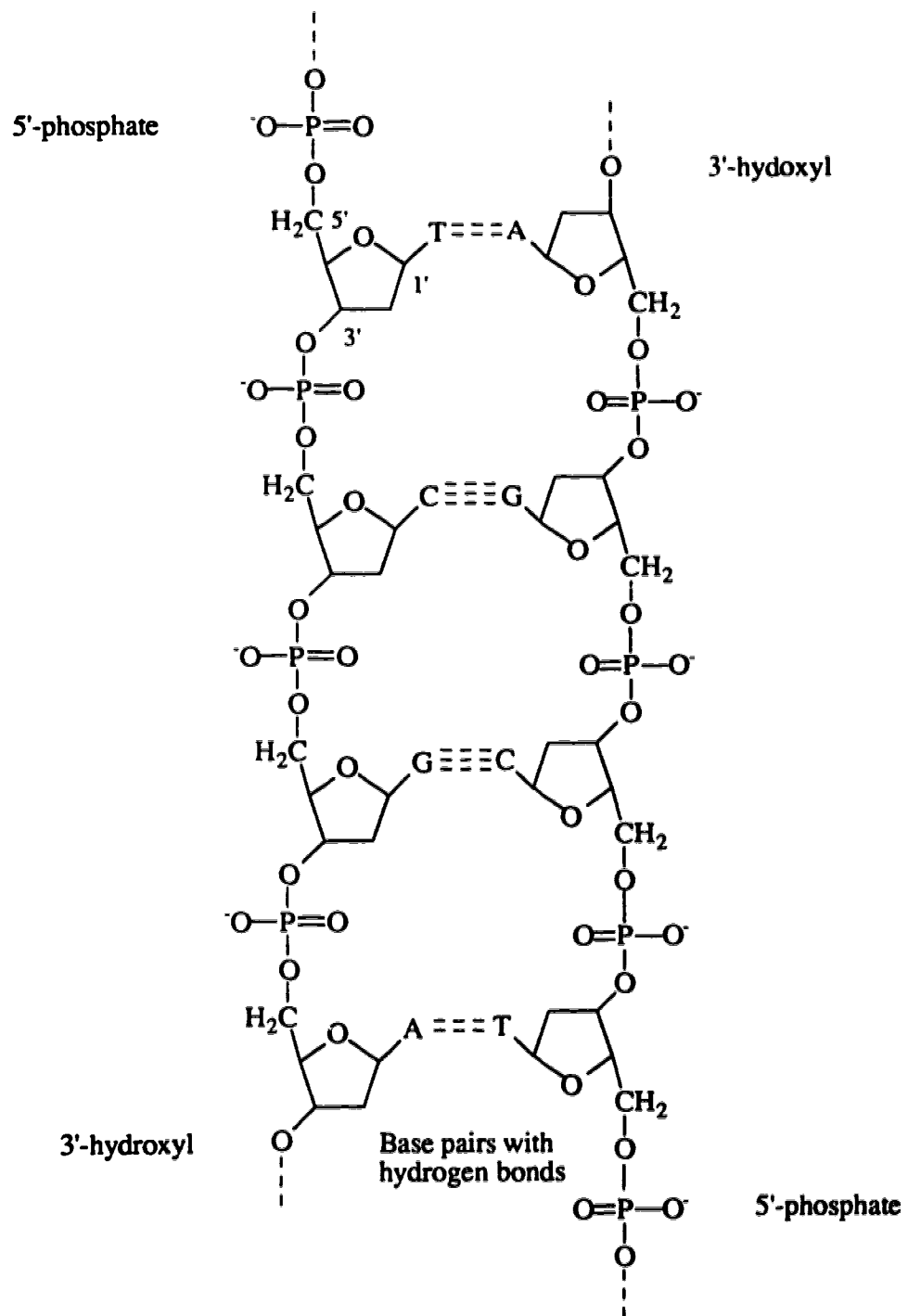


Fig. 1.1. (a) A schematic diagram of DNA structure. A represents Adenine, C represents Cytosine, G represents Guanine and T represents Thymine.

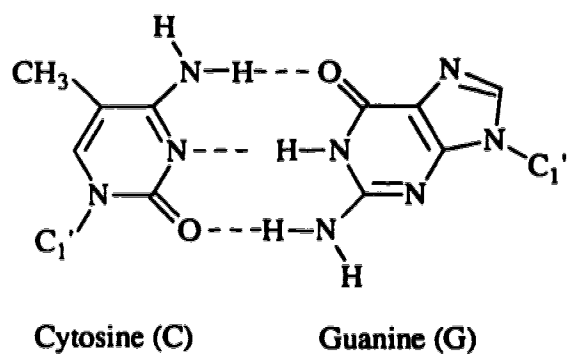
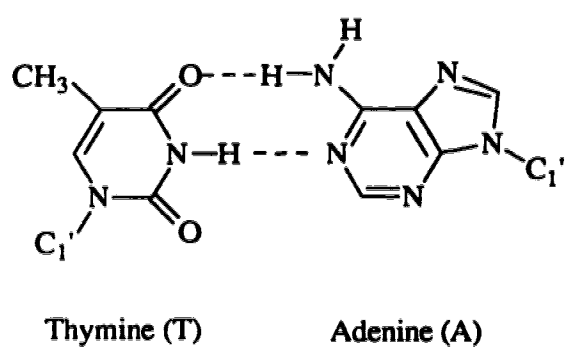


Fig. 1.1. (b) Molecular structures of A, C, G and T. Base-pairs A-T and C-G are formed through hydrogen bonds shown in dashed lines.

1.2. DNA SEQUENCING

In 1977, two now classical and powerful methods were reported for DNA sequencing: the chemical degradation method of Maxam and Gilbert [3], and the chain termination method of Sanger and Coulson [4]. These methods convert the order of nucleotides in the DNA chain, which can not be read directly, into a chain-length information, which can be determined experimentally. In Sanger's method, for example, a primer (a short, single-stranded oligonucleotide) is associated with a template (a single-stranded DNA with unknown sequence). This template-primer solution is divided into four containers, and DNA polymerase (an enzyme) and deoxynucleotide triphosphates (dNTPs) are added to each container. Finally, dideoxyadenosine triphosphate is added to the first container, dideoxycytidine triphosphate is added to the second container, dideoxyguanosine triphosphate is added to the third container, and dideoxythymidine triphosphate is added to the last container. In the absence of the dideoxynucleotide triphosphates, the polymerase catalyzes the extension of the primer, producing a DNA strand that is complementary to the template. In the presence of a small amount of a dideoxynucleotide, the chain extension will occasionally be halted upon incorporation of that dideoxynucleotide. This reaction is known as the chain extension/termination reaction or simply the chain termination reaction. The product of the reaction in the first tube is a set of DNA fragments that have an identical origin (primer) and extend to various lengths, always ending in dideoxyadenosine. Similarly, the products of the other reactions are sets of DNA fragments that end in the other nucleotides. Fig. 1.2 shows a schematic diagram of sample generation by Sanger's method.

Sanger incorporated a small amount of radioactive nucleotide in the sequencing reaction mixture. The radioactive products of the sequencing mixture were separated by electrophoresis in adjacent lanes of a slab polyacrylamide gel. The gel separated the fragments based on size: mobility decreases monotonically with fragment length. After

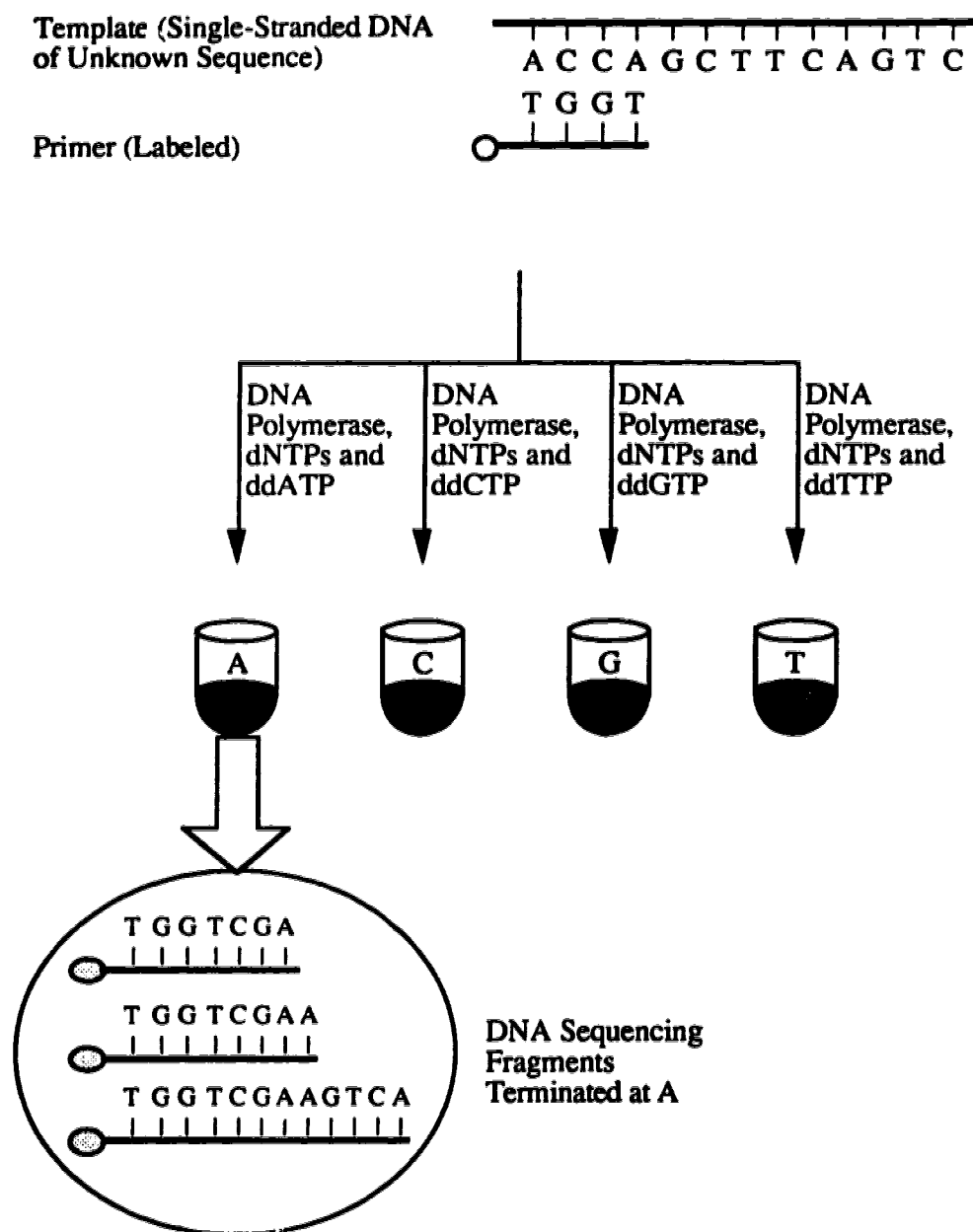


Fig. 1.2. A schematic of sample generation by Sanger's method. For simplicity, only A-terminated sample is shown in detail.

the electrophoresis was completed. x-ray film was placed in contact with the gel, and the location of the DNA fragments were detected by autoradiography. Finally, the sequence was interpreted from the pattern of alternating bands in the lanes corresponding to the terminal base of the fragment (Fig. 1.3).

After Sanger's technology became widely available, DNA sequencing became a routine, albeit tedious, technique in the biomedical laboratory. Sanger's method requires adjacent lanes. The order of the terminating nucleotides or DNA sequence is determined roughly 8-16 hours for the electrophoretic separation, 8-16 hours for autoradiography, and perhaps another 4 hours for visual inspection of the autoradiogram. In addition, Sanger's method requires use of radionuclides; the expense of disposal of low-level radioactive waste is rapidly becoming a significant item in the sequencing budget.

Advances in sequencing technology occurred in 1986-1988 when Smith *et al.* [5] in Hood's laboratory, Ansorge *et al.* [6], Prober *et al.* [7] at DuPont, and Kambara *et al.* [8] at Hitachi reported DNA sequencing that replaced the radioactive labels and autoradiography in Sanger's method with fluorescent labels and laser-based detection. By use of fluorescence detection, the difficulties associated with autoradiography are eliminated. Furthermore, by use of computerized data collection, the interpretation of the electropherogram to obtain sequence information may be automated. As a result, these fluorescence instruments determined DNA sequence within an 8- to 12-hour period after introduction of the DNA sequencing sample into the slab gel. Gels of 0.2- to 0.5-mm thickness are employed in these slab-gel sequencers. Joule heating limits the applied electric field to about 50 V/cm; overall sequencing rates are about 1000 bases per hour for each slab gel. Sequence may be determined by use of computer algorithms to about 450 bases, which is limited by the signal/noise ratio of the data and the resolution of the gel. Ansorge's group has recently demonstrated a sequencing length of more than 1000 bases on a 50-cm separation distance. The sequencing rate is about 100 bases/hour [9].

Smith's, Ansorge's, Prober's, and Kambara's fluorescence sequencers have been

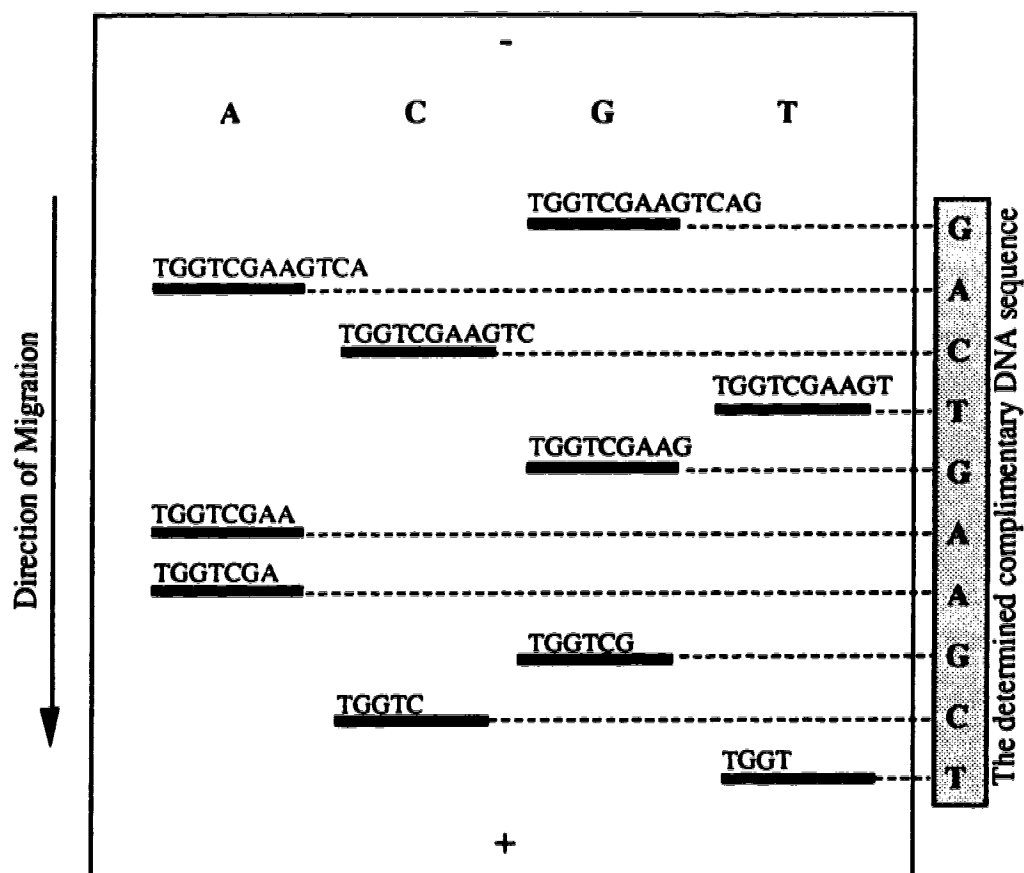


Fig. 1.3. The conventional DNA sequencing by gel electrophoresis in the slab format. Single-stranded DNA fragments terminated with A, C, G and T are separated in four by the relative distance the fragments migrate. Each thick line is the electrophoretic band with the corresponding single-stranded DNA fragment printed above. The complementary DNA sequence is determined at the right side of the slab.

commercialized by Applied Biosystems, Pharmacia, DuPont, and Hitachi, respectively; as an example of typical performance, the Applied Biosystems model 373A instrument runs 24 lanes simultaneously on a slab gel to produce sequencing rates of 50 bases per hour per lane or 1200 bases per hour per slab. Similar sequencing rates are produced by other instruments.

Increased sequencing rates are produced by operating the gels at high electric field strength. Unfortunately, the finite resistance of the separation buffer leads to unacceptable heating of conventional 0.2- to 0.5-mm-thick gels at electric field strengths much greater than 100 V/cm. Gel-filled capillaries have attracted interest because their high surface-to-volume ratio provides excellent heat transport properties, allowing use of very high electric field strengths.

1.3. DNA SEQUENCING APPROACHES

Many approaches have come into being with the aim of fast, convenient and efficient sequence determination. Particular attention has been paid to those approaches where fluorescent labels are incorporated.

1.3.1. ONE-DYE PRIMER (PHARMACIA)

The Pharmacia Automatic Laser Fluorescent (ALF) DNA sequencer (as well as a similar instrument developed by Hitachi) uses the one-dye-four-lane approach developed by Ansorge *et al.* [6]. This approach, by replacing the radioactive nucleotide with a fluorescently labeled-primer, is very similar to traditional Sanger's method: it uses only one fluorophore and requires four lanes to complete DNA sequence determination. It is simple in terms of sample preparation, it requires no differential mobility adjustment, but it needs to compare the data from four adjacent lanes and is low in terms of sequencing

throughput.

It is interesting to note that in the ALF DNA sequencer, a single laser beam is directed through the gel across its width; a set of 40 identical, discrete photodetectors, one per lane, is used [10]. This single-laser and 40-detector design allows for the simultaneous recording of sequences from 10 samples. Because the detection is done in parallel, the run time is relatively short (about 6 hours).

1.3.2. FOUR DYE-PRIMERS (ABI)

Applied Biosystems Inc. (ABI) has developed a set of four dye-labeled primers based on Smith's method [5]. The four fluorescent dyes are FAM, JOE, TAMRA and ROX (nicknames for fluorescein, 2',7'-dimethoxy-4',5'-dichloro-fluorescein, tetramethylrhodamine, and rhodamine χ , respectively). Four sequencing reactions are required to generate DNA sequencing products, each of them is labeled with one specific dye. These products are combined to form a sequencing sample and subjected to electrophoretic separation in one lane. Sequence information is encoded in the color of the dyes and the relative electrophoretic mobility of the DNA fragments.

Since four dyes are used in this approach, four separate sequencing reactions are required to generate a sequencing sample. However, the mobility might be different for the same DNA fragments labeled with different dyes and the problem of this mobility shift needs to be taken into account [11]. Because the separation is conducted in a single lane, data comparison from adjacent lanes is not necessary, and the sequencing throughput is increased. Twenty four samples may be sequenced simultaneously.

1.3.3. FOUR DYE-TERMINATORS (DUPONT AND ABI)

DuPont first introduced a dye-terminator approach for DNA sequence

determination [7]. A family of four succinylfluorescein dyes are used to label the dideoxy terminators. Fig. 1.4 shows these dye-labeled terminators. These dyes are very similar in structure and as a result, their absorbance and emission spectra are closely spaced. The closely spaced absorbance spectra allows a single laser line (488 nm from an argon ion laser) to excite all of the four dyes. The excited fluorescence is distributed into two spectral channels. The ratio of the signals from these two channels is used to determine the corresponding terminal bases.

Several advantages are inherent in this approach. First, the DNA sequencing sample is prepared in one tube instead of four tubes and the electrophoretic separation is also conducted in one lane. One has the freedom to choose primers. Second, in this approach, only true dideoxy termination result in detectable fluorescent signal. Chain synthesis may be terminated by deoxynucleotides, but this will not be detected since they do not fluoresce. Third, the instrumentation is simple, because only one laser line is for excitation and two spectral channels are for detection.

The major and fatal drawback of this approach is its poor accuracy. Background noise, anomalous migration patterns and some other sources lead to errors in sequence determination. Because of its low accuracy, this approach has been phased out by the company. But the idea of dye-labeled terminators is still alive and has proved to be valuable.

ABI has produced two sets of dye-terminators for DNA sequencing. Its second generation of dye-terminators [12] is called the PRISM system. Instead of using very similar dyes, the PRISM system involves a family of fluorescein derivatives that have relatively widely spaced absorbance and emission spectra. Fig. 1.5 shows the structure of the dyes along with the absorbance and emission maxima. Two laser lines from an argon ion laser, 488 and 514.5 nm, are required for fluorescence excitation from the four dyes and four spectral channels are used for fluorescence detection. The instrument is more complicated than that marketed by DuPont, but DNA sequence determination is more

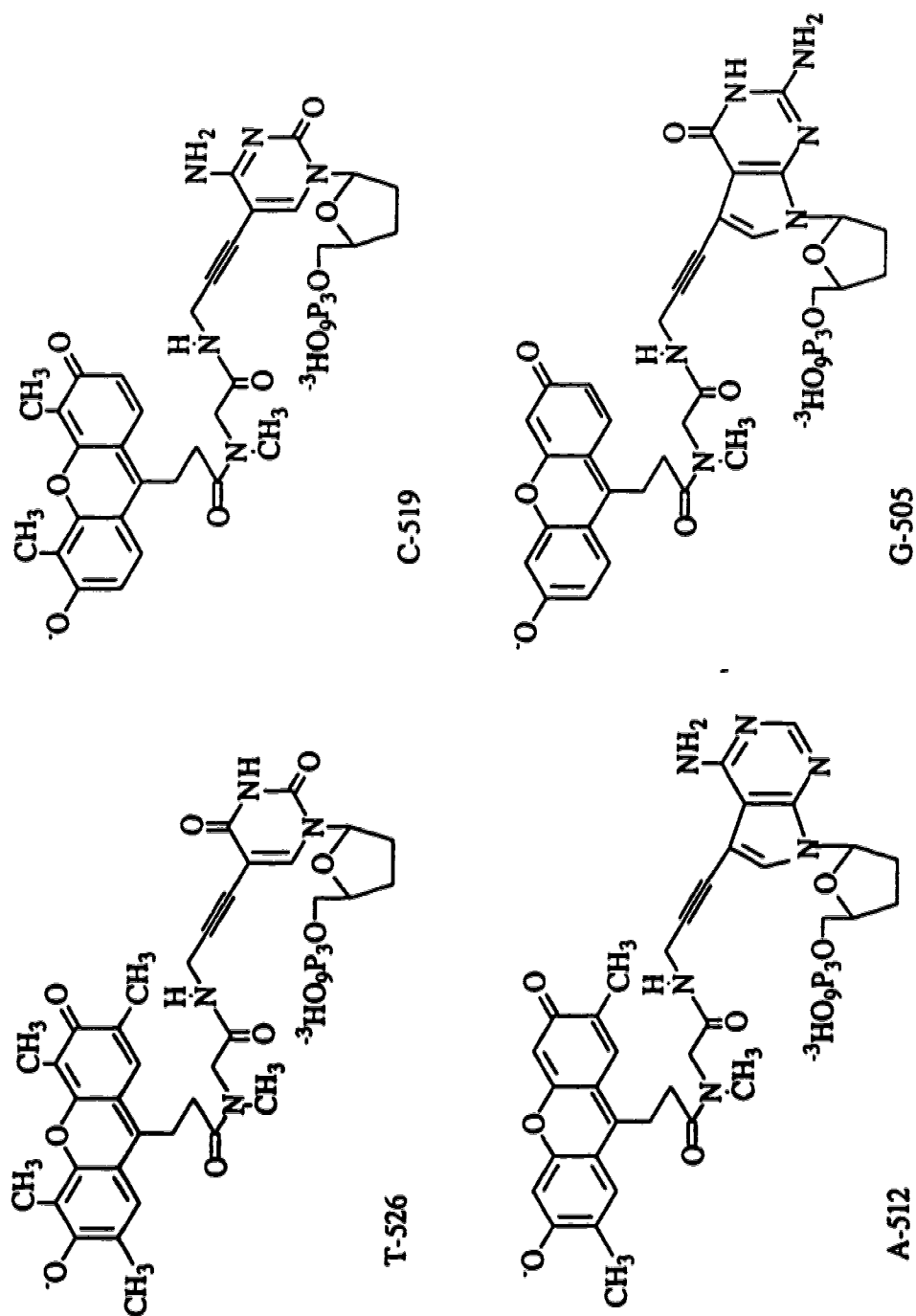
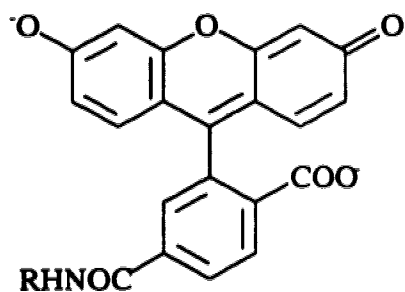
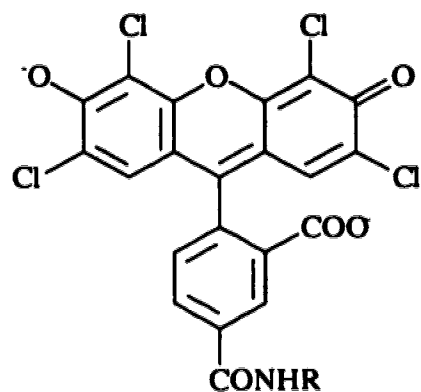


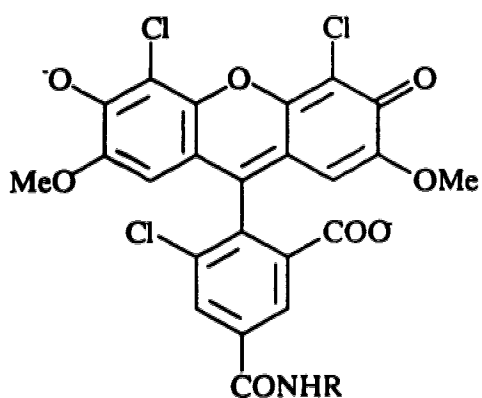
Fig. 1.4. Structures of DuPont's dye-terminators. The numbers included in the names indicate the wavelength maxima of fluorescence in nm.



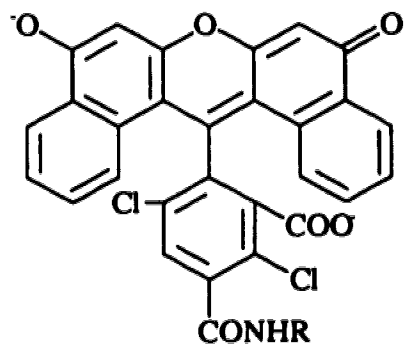
ddT-6FAM
 Abs max = 494 nm
 Em max = 520 nm



ddC-5ZOE
 Abs max = 520 nm
 Em max = 540 nm



ddA-LOU2
 Abs max = 538 nm
 Em max = 550 nm



ddG-NAN2
 Abs max = 556 nm
 Em max = 585 nm

Fig. 1.5. Four fluorescein dye-terminators used in ABI's PRISM system (R represents the corresponding nucleotide).

accurate.

1.3.4. DYE-DEOXY OR INTERNAL LABEL

Ansorge's group has developed an approach in which the deoxynucleotide is labeled with a fluorescent dye [9]. After annealing of primer to template, they labeled the DNA fragments by extending the chain with fluorescein-15-*dATP for a short distance and then terminated the chain with a deoxy- and dideoxy-nucleotide mixture. This approach allows multiple labeling on a single chain provided that the sequencing bands are not distorted. Usually, preparation of dye-labeled primers is expensive and time-consuming; the use of dye-labeled deoxynucleotide facilitates sample preparation for DNA sequencing. Ansorge's group has demonstrated the reading length that extended to 1000 bases on gels with 50-cm separation length [13]. Because only one dye-labeled deoxynucleotide is commercially available now, the DNA sequencing fragments need four lanes for separation.

1.3.5. ONE-DYE PEAK-HEIGHT

In 1990, Richardson and Tabor [14] and, independently, Ansorge [15] reported a sequencing approach based on a single dye-labeled primer. By variation of the amount of dideoxynucleotide (in the ratio of 8:4:2:1, for example) in the reaction mixture, each base is identified with a particular peak height during separation in a single lane of polyacrylamide gel. This approach relies on uniform labeling of the reaction products through use of manganese/T7 polymerase reaction. This one-label four-peak-height approach is very efficient in sample preparation and separation, but it is not very accurate. Several sources can be identified that lead to errors in sequence determination [16]. Peak height variation associated with each ddNTP is about 25%. There are chances of missing

peaks when small peaks are sandwiched between two large peaks. Systematic errors exist due to the presence of ghost peaks. As a result, this technique suffers from low accuracy, typically 90% for fragments shorter than 250 bases.

1.3.6. TWO-DYE, PEAK-HEIGHT

To improve the accuracy of the sequence determination and yet retain the simplicity of the peak-height approach, Dovichi and co-workers have reported a modified version—a two-label, peak-height encoding approach [17,18]. Two fluorescently labeled primers are employed in separate sequencing reactions. For example, one sequencing reaction uses FAM-labeled primer with ddATP and ddCTP; the concentrations of ddATP and ddCTP are adjusted to produce a 3:1 ratio in the fluorescence intensity. The other sequencing reaction uses a TAMRA-labeled primer with ddGTP and ddTTP; the concentrations of ddGTP and ddTTP are also adjusted to produce a 3:1 ratio in fluorescence intensity. The pooled reaction products are separated in a single capillary (or lane). The fluorescence is detected in two spectral channels. Both color and intensity of fluorescence are used to identify the terminal bases.

This approach is a compromise between one-label, four-peak-height and four-label techniques—it requires two separate sample preparations and two spectral channels for detection. Accuracy as high as 99.5% has been achieved by this approach for the first 400 bases.

1.3.7. TWO-DYE BINARY AND ONE-LABEL THREE-HEIGHTS

Two other variations to peak-height approach are binary coding [19] and one-label three-height [20] approaches. In the binary coding approach, two labels are used in a binary fashion and two spectral channels are used for detection. For example, A is labeled

with both FAM and JOE, C is labeled with only FAM, G is labeled with only JOE, and T is not labeled. If 1 denotes the presence of signal in one channel and 0 denotes the absence of signal, the DNA fragments are binary coded: A is {1, 1}, C is {1, 0}, G is {0, 1} and T is {0, 0}. The {0, 0} code simply leaves a gap in the DNA sequence.

The one-label, three-height approach is quite straightforward: the concentrations of three dideoxynucleotides are adjusted to give a peak height in the ratio of 4:2:1, with the other dideoxynucleotide absent in the sequencing sample.

In both approaches, one of the fragments is unlabeled and is detected as a gap in the DNA sequence. DNA sequencing should be conducted twice to generate complete sequence information (for example, A, C, and G are labeled in the first sequence determination and T, G, and C in the second). The argument is that the two data files can be combined to make a more confident sequence determination. But compressions in the sequence as a result of migration anomaly of the DNA fragments lead to serious difficulties with use of unlabeled fragments [21].

1.3.8. CYCLE SEQUENCING

Cycle sequencing is related to the polymerase chain reaction (PCR).

PCR is a procedure by which a certain segment of DNA can be exponentially amplified by a factor of 1 million or more through many cycles with each cycle doubling the amount [22]. In PCR, two primers of known sequence, which flank the segment, initiate double-strand DNA synthesis in the presence of nucleotides and polymerase. Essentially, there are three steps in each cycle of PCR: 1) high-temperature denaturation or separation of double-stranded DNA into its component single strands; 2) primer annealing or hybridization of primers to their complementary sequences on the single strands; and 3) chain extension or addition of nucleotides to the new chains by polymerase. Many cycles are performed in this way, exponentially amplifying the desired

segments (Fig. 1.6a). The use of thermostable polymerase such as *Taq* DNA Polymerase allows PCR be carried out automatically without the addition of fresh polymerase after each cycle.

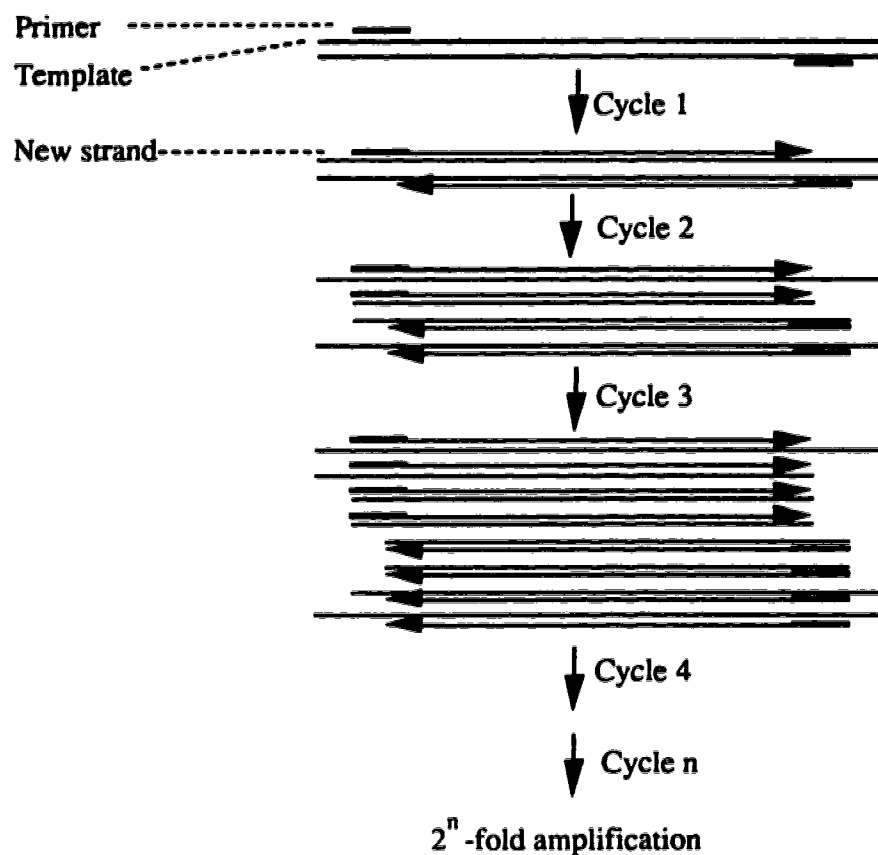
DNA sequencing fragments can be produced through thermal cycling with the aid of thermostable polymerase. This procedure is called cycle sequencing (Fig. 1.6b) [23, 24]. In cycle sequencing, only a single primer is used to linearly amplify a segment of template DNA, and dideoxynucleotide triphosphates are used to terminate each newly synthesized chain. After many cycles, an amplified nested set of fragments is produced just like that produced by conventional Sanger's method. Cycle sequencing is a very interesting approach by which a small amount (on the order of nanograms) of template is required to produce ample quantities of sequencing products. In principle, DNA may be sequenced directly from single phage plaques, avoiding growth of a large amount of template. By cycle sequencing, sequence information may be obtained directly and efficiently from double strand template. Large-scale sequencing efforts may benefit from cycle sequencing.

1.3.9. PRIMER WALKING BY CONTIGUOUS HEXAMERS

In DNA sequencing, gel resolution limits the sequencing length to several hundred to one thousand bases. Sequencing of long stretches of DNA is usually accomplished through successive generation of many short sequences. Two general approaches are currently applied: random (or shotgun) and directed (walking) [25]. In large-scale DNA sequencing, the two approaches may be combined to gain flexibility [11].

The shotgun strategy generates many short sequences (called subclones) at random (Fig. 1.7a). To put them in order, probable overlaps of such sequences should be identified. To complete the sequence, a large number of extra (or redundant) subclones

a. PCR



b. Cycle sequencing

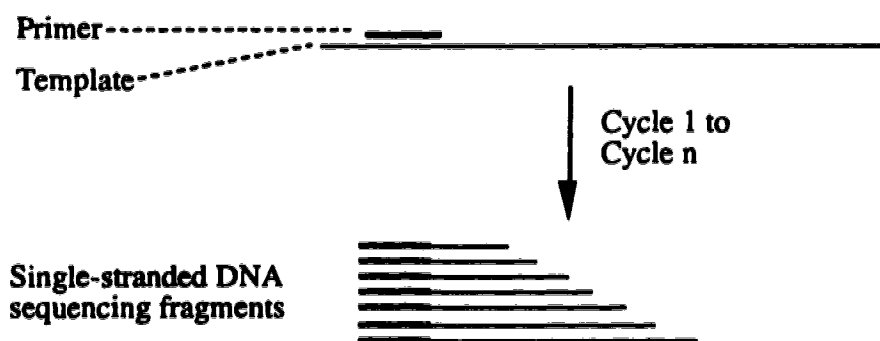
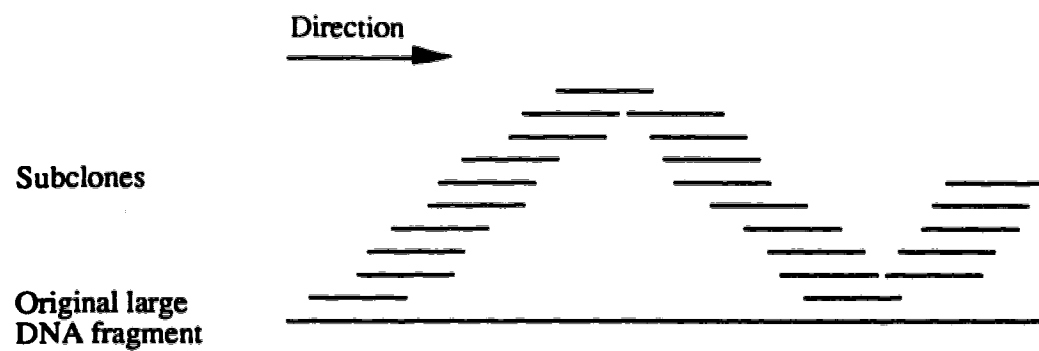


Fig. 1.6. Schematic diagrams of (a) PCR and (b) cycle sequencing. In cycle sequencing, a single primer is used for linear amplification of the DNA sequencing fragments.

a. Shotgun (random)



b. Primer walking (directed)

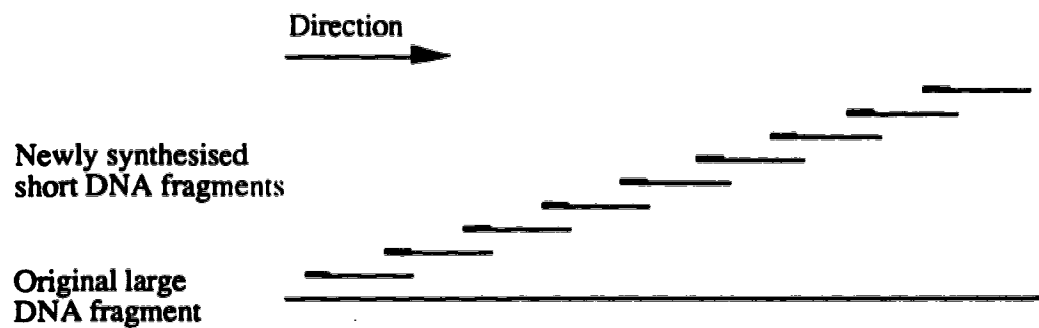


Fig. 1. 7. Schematic diagrams of (a) shotgun strategy and (b) primer walking strategy.

need to be sequenced.

On the other hand, the primer walking strategy (Fig. 1.7b) obtains the sequence information in a fairly ordered fashion: sequencing continues from the beginning to the end along a DNA chain. The just-acquired sequence information is used to synthesize a primer (about 15 to 20 bases long) for subsequent new sequence determination. This approach should be much more efficient compared to the shotgun strategy because there is no need for subcloning and multiple template preparation. Primer walking offers low redundancy and the sequence assembly is quite straightforward. The limitation of the walking strategy is the requirement of synthesis of very large numbers of primers; this synthesis is slow and expensive.

Recently, Kieleczawa, Dunn and Studier at the Brookhaven National Laboratory announced a major advance in the primer walking approach to DNA sequencing [26]. In their approach, three hexamer strings can be used to form 18-nucleotide primers when the template is coated with a single-strand binding protein and the temperature is controlled at 0 °C. In primer walking, it is better to use primers that can be stored in a library rather than to synthesize a primer for each new use. A complete library of 4^6 or 4096 possible hexamers could conveniently supply the primers needed for large scale DNA sequencing.

1.4. CAPILLARY FREE-ZONE ELECTROPHORESIS

Electrophoresis has proven to be very useful in separating ionic analytes, particularly those macro-molecules of biomedical importance [27]. In the early days, however, the separation efficiency of electrophoresis was limited by convection caused by Joule heating. Consequently, traditional electrophoresis is usually performed on a solid support such as paper, or in an anti-convective media, such as polyacrylamide or agarose gel, either in slab or tube format.

In the early 80's, Jorgenson and Lukacs investigated and began to popularize

capillary electrophoresis [28]. Since that time, the technique has attracted significant attention in the analytical community. The small dimension of the capillaries, typically 10 to 100 μm inner diameter (ID), minimizes convection [29]. Also, this small dimension greatly reduces current, and hence heat, generated by electrophoresis. The large surface-to-volume ratio provides good heat dissipation, allowing very high electric fields to be applied so as to achieve both high efficiency and high speed separation.

In free zone electrophoresis (usually called zone electrophoresis for simplicity), for example, the capillary is filled with buffer solution and placed between two buffer vessels. Two electrodes are used to achieve electrical contact (Fig. 1. 8). The analytes migrate as zones along the capillary and the separation is based on the differential movement of the zones under the influence of the electric field. Fundamentally, two processes are associated with capillary zone electrophoresis: electrophoresis of the analytes and the electroosmosis of the bulk solution. Electroosmosis or electroendoosmosis is an important phenomenon in understanding capillary electrophoresis. Under aqueous conditions, the silanol groups on the surface of fused silica capillary tend to deprotonate, resulting in a negatively charged surface. Counterions, which build up near the surface to maintain the charge balance, form the double layer. The cations in the diffuse double layer are attracted to the negative end of the capillary upon the application of an electric field, dragging the bulk solution toward the negative end. Normally, the effect of electroosmosis is so strong that all analyte species, whether they are positively charged, negatively charged, or neutral, will flow from the anode to the negative end. For neutral species, the migration rate is identical to that of electroosmotic flow (v_o)

$$v_o = \mu_o \frac{V}{L} \quad (1-1)$$

where μ_o is the coefficient of electroosmotic flow, V is applied voltage, and L is the

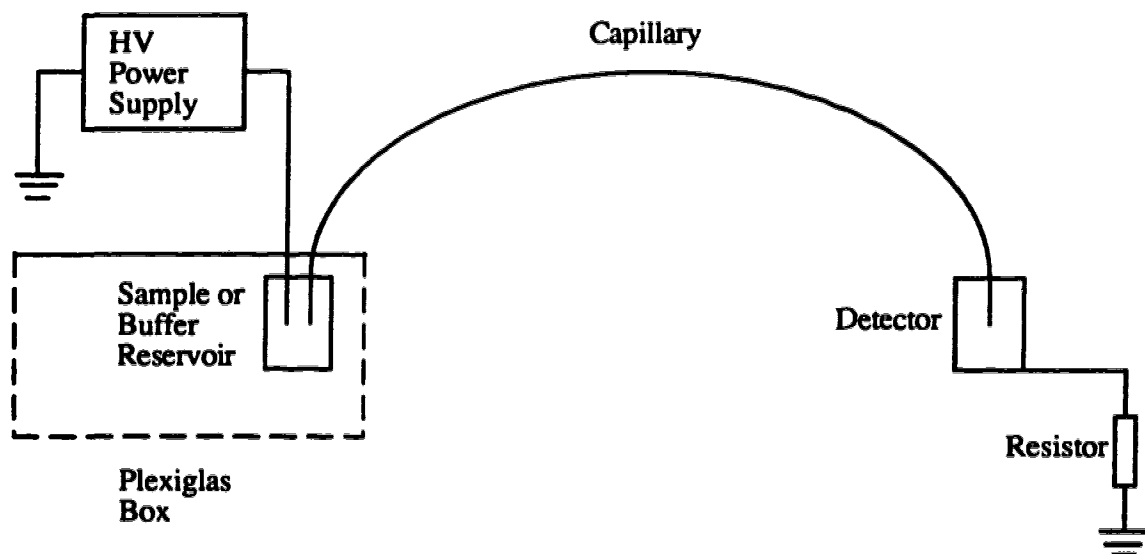


Fig. 1.8. Schematic of a capillary electrophoresis system (post-column detector).

length of the capillary. For charged species, the migration rate, v , is the superimposition of electrophoretic flow rate and electroosmotic flow rate

$$v = v_e + v_o = (\mu_e + \mu_o) \frac{V}{L} \quad (1-2)$$

where μ_e is the electrophoretic mobility of the ionic species.

The time, t , required for an analyte to migrate the entire length of the capillary is given by

$$t = \frac{L}{v} = \frac{L^2}{(\mu_e + \mu_o)V} \quad (1-3)$$

For a well designed zone electrophoresis system where mass diffusion is the dominant factor for zone broadening, the spatial variance, σ^2 , of a zone after a time, t , can be derived from the Stokes-Einstein equation:

$$\sigma^2 = 2Dt = \frac{2DL^2}{(\mu_e + \mu_o)V} \quad (1-4)$$

where D is the diffusion coefficient of the analyte in the zone.

By analogy to chromatography, the separation efficiency of an electrophoretic system can be expressed in terms of theoretical plates

$$N = \frac{L^2}{\sigma^2} = \frac{(\mu_e + \mu_o)V}{2D} \quad (1-5)$$

and the resolution, R_s , of two zones in capillary electrophoresis can be expressed as

$$R_S = \frac{N^{1/2}}{4} \cdot \frac{\delta v}{\bar{v}} = \frac{N^{1/2}}{4} \cdot \frac{\mu_{e1} - \mu_{e2}}{\mu_e + \mu_o} \quad (1-6)$$

where $\frac{\delta v}{\bar{v}}$ is the relative velocity difference between the two zones, μ_{e1} and μ_{e2} are the electrophoretic mobility of the two analytes, and μ_e is their average electrophoretic mobility.

By combining Equations (5) and (6), we can obtain

$$R_S = 0.18(\mu_{e1} - \mu_{e2}) \left(\frac{V}{D(\mu_e + \mu_o)} \right)^{1/2} \quad (1-7)$$

Equations (3), (5) and (7) indicate the major advantages inherent in capillary electrophoresis: with the application of high voltage, we can achieve short separation time, high separation efficiency, and high resolution. All these advantages arise from the large surface-to-volume ratio of the capillary.

It should be noted that the separation efficiency is generally higher for capillary electrophoresis than for liquid chromatography for at least two reasons. First, the separation for capillary zone electrophoresis is based on charge-to-size ratio. Analytes are separated as zones that migrate along with the bulk flow. High potential can drive the electrophoresis at high speed and with high efficiency; for liquid chromatography, the separation is based on partitioning of species between the mobile phase and stationary phase through mass diffusion. The mass diffusion is usually a slow process, particularly for large molecules. As a result, the mass transport rate in the mobile phase, or separation speed, is limited and zone broadening can be substantial. Second, in capillary zone electrophoresis, the bulk flow is electrically “pumped” by electroosmotic flow, generating a flat flow profile. This flat flow profile is beneficial because it does not directly contribute to the dispersion of analyte zones. In liquid chromatography, the flow

is externally pumped. The flow profile is laminar or parabolic, generating an additional contribution to band broadening.

1.5. CAPILLARY GEL ELECTROPHORESIS

As mentioned above, free zone electrophoretic separation is based on the charge-to-radius ratio of the analyte. Unfortunately, in DNA, this ratio remains roughly constant as additional bases are added to a DNA chain. In contrast to zone electrophoresis, gel electrophoresis produces separation that is based on the size of the analyte. The gel medium, usually polyacrylamide or agarose, acts as a "molecular sieve". As DNA fragments migrate through the gel network, they become retarded, with large fragments retarded more than small ones. DNA sequencing is almost exclusively conducted in crosslinked polyacrylamide gel, the composition of which is expressed in terms of %T and %C:

$$\%T = \frac{W_{\text{Acrylamide}}(\text{g}) + W_{\text{Bis}}(\text{g})}{V(\text{ml})} \times 100\%$$

$$\%C = \frac{W_{\text{Bis}}(\text{g})}{W_{\text{Acrylamide}}(\text{g}) + W_{\text{Bis}}(\text{g})} \times 100\%$$

where $W_{\text{Acrylamide}}$ and W_{Bis} are the weights of acrylamide and N,N'-methylene bisacrylamide or Bis in grams; V is the total volume of the solution in ml. The extent of such retardation is a function of the gel composition and can be expressed by a Ferguson plot [30]. To some extent, the plot of log (mobility) versus %T at constant %C is a linear relationship as follows:

$$\log \mu = \log \mu_0 + k (\%T) \quad (1-8)$$

where μ_0 is the mobility in free solution (0 %T), k is the retardation coefficient and %T is the acrylamide concentration. Due to the retardation effect in gel electrophoresis, there often is a linear relationship between fragment size (bases) and migration time.

Conventional crosslinked gel is comprised of two parts: a solid polymer network and liquid medium [31]. The network prevents liquid from escaping, and the liquid helps to maintain the network. In capillary gel electrophoresis, the gel is not important as an anti-convective medium. The polymer network can be conventional crosslinked polyacrylamide, or the polymer can be simply a linear polymer such as polyacrylamide without crosslinker. The former is also called a chemical gel, or crosslinked gel, whereas the latter is also called a physical gel, non-crosslinked gel or entangled polymer. Although there is a radical difference in the gel structure, the separation mechanism of the gels seems identical. While non-crosslinked gel at low concentration (lower than 10 %T) is not stable in thick slab format [32], it is well behaved in capillary format for DNA sequencing [18]. Gel electrophoresis in the capillary format offers several advantages over conventional slab gel electrophoresis, including high speed, high resolution, and ease of automation. These advantages are very valuable for both routine separation and large-scale DNA sequencing such as the Human Genome Project.

In addition to the applied voltage, the speed of DNA sequencing can also be controlled through experimental conditions such as the choice of the gel matrix (polyacrylamide, Long Ranger [33] and Hydrolink [13]), the concentration of the gel, the composition of the gel (e.g. the addition of formamide to the gel [16]) and the temperature of the gel [34].

1.6. LASER-INDUCED FLUORESCENCE DETECTION

One big issue associated with capillary electrophoresis is detection. To produce reproducible and efficient separations, the analyte concentration must be much lower than

the ion concentration of the separation buffer. Besides, a capillary with 50 μm id. and 50 cm in length has a volume of less than 1 μl . Therefore the sample volume injected should be on the order of a nanoliter or less. Currently, high-sensitivity detectors used for capillary electrophoresis include those based on mass spectrometry [35], electrochemistry [36], laser-induced thermo-optical absorbance [37], and laser-induced fluorescence [38]. Laser-induced fluorescence is particularly interesting because of its ultimate detection sensitivity—counting the molecules of analyte within a sample [39].

Fluorescence intensity (I) of a molecule is related to some physical parameters as follows :

$$I = QI_0[1 - 10^{-abc}] \quad (1-9)$$

where Q is the quantum yield of the molecule, I_0 is the incident power of the exciting beam, a is the molar absorptivity of the molecule, b is the sample path-length, and c is the molar concentration of the sample. When the sample concentration is low and the path-length is short, Equation (9) reduces to

$$I = QI_0abc \quad (1-10)$$

This equation shows a linear relationship between fluorescence intensity and sample concentration and this relationship is only obeyed for dilute solutions. At high concentration, self-absorption and concentration quenching lead to some deviations.

Because of the fact that fluorescence emits in all directions but can only be viewed within a limited aperture, Equation 1-10 can be written as

$$I = KQI_0abc \quad (1-11)$$

to include an instrumental factor K , where $K < 1$. Since sample concentration (c), and sample path-length (b), are necessarily low in capillary format, what we can do is to choose a molecule that has high quantum yield (Q) and high molar absorptivity (a) at a convenient excitation wavelength, increase the source intensity (I_0), and increase the fluorescence collection efficiency (related to K).

Since most analytes do not fluoresce, labeling the sample with fluorescent dyes is often necessary. The samples most frequently used in this respect can be different biochemical molecules such as amino acids, peptides, proteins, oligosaccharides, oligonucleotides and DNA molecules. One of the most widely used labels is fluorescein. Fig. 1.9 and Fig. 1.10 are the normalized absorbance and emission spectra of this molecule, respectively. The laser line, 488 nm, from an argon ion laser is also indicated. The 488 nm line from an argon ion laser is well suited to the excitation of fluorescein as can be seen from the absorbance spectrum. Fluorescein labeled analytes have high molar absorptivities and fluorescence quantum yields, $\sim 80,000 \text{ l}\cdot\mu\text{mole}^{-1}\cdot\text{cm}^{-1}$, and ~ 0.5 , respectively [40]. It should be pointed out that in DNA sequence determination using four dye systems, some excitation wavelengths and emission passbands are chosen away from the spectral maxima because of the limited choice of laser lines and required spectral resolution.

Laser-induced fluorescence is well suited for detection in capillary electrophoresis. The spatial coherence of the laser allows very small volumes to be excited with high intensity. When coupled with efficient collection and detection, minute amounts of analyte may be detected. Detection limits in fluorescence measurements are usually determined by shot-noise in the background signal. Parker [41] listed four sources of background signal in fluorescence: fluorescence from solvent impurities, fluorescence from cuvette windows, Raman and Rayleigh scatter from solvent, and light scatter at the cuvette-sample interface. Detector dark current is a fifth source of background. To produce excellent detection limits, one must minimize each of these sources of

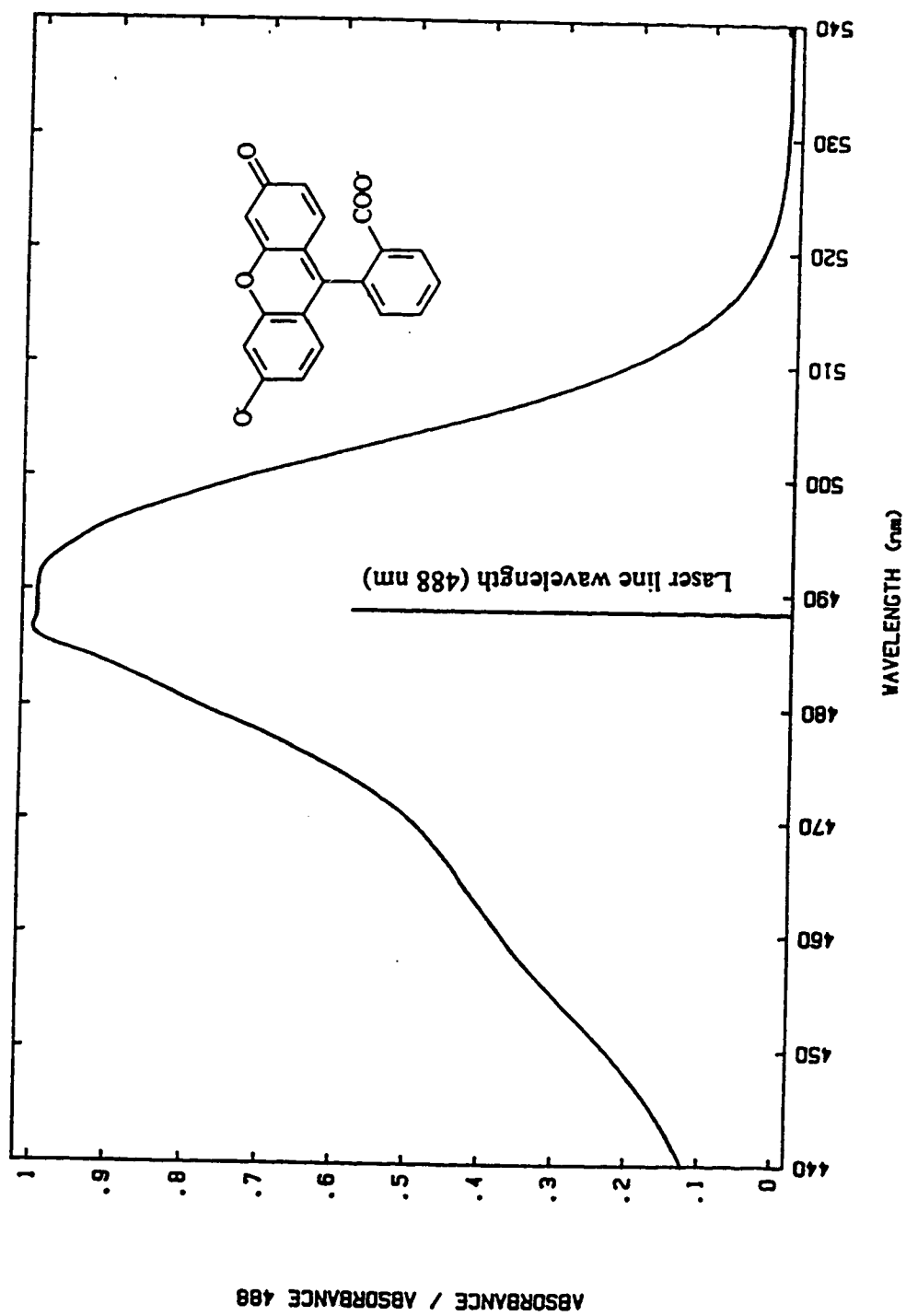


Fig. 1.9. The normalized absorbance spectrum of fluorescein. The measurement was performed on a Hewlett-Packard diode array spectrophotometer, model 8451 A. The fluorescein solution is 3.1×10^{-6} M in 10 mM borate buffer, pH 9.2.

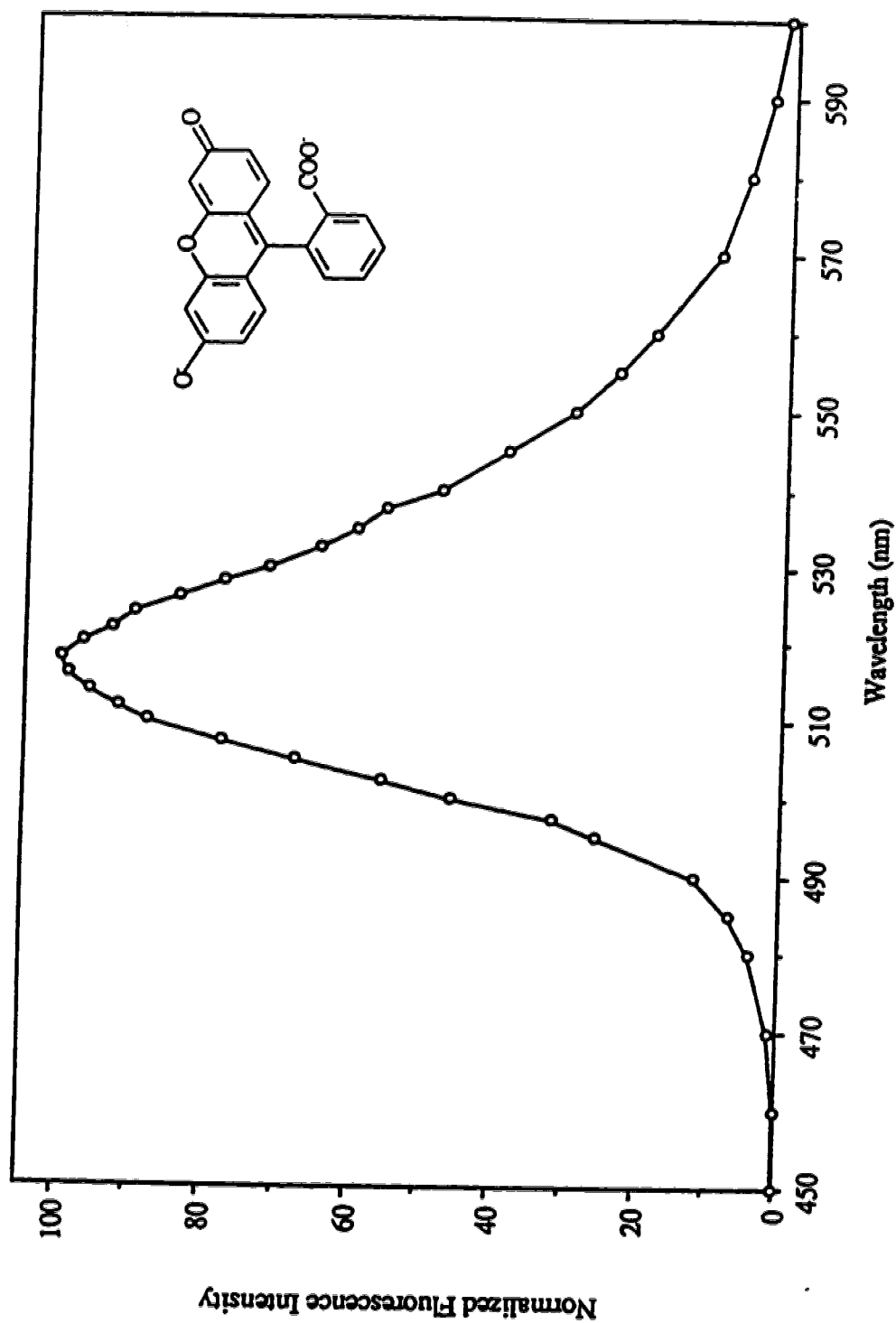


Fig. 1.10. The normalized fluorescence spectrum of fluorescein (uncorrected). The measurement was performed on a Turner Spectrofluorometer, model 430. The fluorescein solution is 3.1×10^{-6} M in 10 mM borate buffer, pH 9.2.

background signal.

Reagent contamination is minimized by careful sample preparation. Solvent Raman scatter may be rejected through judicious spectral filtering. Fluorescence and light scatter from cuvette windows are minimized by careful spatial masking to restrict the field of view of the detector to the illuminated sample stream. Usually, light scatter from the capillary window is the major source of background signal in fluorescence detection in capillary electrophoresis.

A post-column fluorescence detector based on the sheath flow cuvette effectively eliminates the scattered light from capillary interfaces, lowers the background signal to nearly Raman scatter limited [42] or even dark current limited [43]. The sheath flow cuvette is also unique in providing a well defined isolation for multiple sample streams. This cuvette and its performance in capillary electrophoresis is the subject of the remainder of this thesis.

REFERENCES

1. Watson, J.D. and Crick, F.H.C., *Nature*, 1953. **171** : p. 737-738.
2. Watson, J.D. and Crick, F.H.C., *Nature*, 1953. **171** : p. 964-967.
3. Maxam, A.M. and Gilbert, W., *Proc. Natl. Acad. Sci. USA*, 1977. **74** (2): p. 560-564.
4. Sanger, F., Nicklen, S., and Coulson, A.R., *Proc. Natl. Acad. Sci. USA*, 1977. **74** : p. 5463-5467.
5. Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B.H., and Hood, L.E., *Nature*, 1986. **321** : p. 674-679.
6. Ansorge, W., Sproat, B., Stegemann, J., Schwager, C., and Zenke, M., *Nucleic Acids Research*, 1987. **15** (11): p. 4593-4602.
7. Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A., and Baumeister, K., *Science*, 1987. **238** : p. 336-341.
8. Kambara, H., Nishikawa, T., Katayama, T., and Yamaguchi, T., *Biotechnology*, 1988. **6**: p. 816-821.
9. Voss, H., Wiemann, S., Wirkner, U., Schwager, C., Zimmermann, J., Stegemann, J., Erfle, H., Hewitt, N.A., Rupp, T., and Ansorger, W., *Methods in Molecular and Cellular Biology*, 1992. **3**: p. 153-155.
10. Freeman, M., Baehler, C., and Spotts, S., *Bio/Technology*, 1990. **8**: p. 147-148.
11. Hawkins, T.L., Du, Z., Halloran, N.D., and Wilson, R.K., *Electrophoresis*, 1992. **13** : p. 552-559.
12. Lee, L.G., Connell, C.R., Woo, S.L., D., C.R., McArdle, B.F., Fuller, C.W., Halloran, N.D., and Wilson, R.K., *Nucleic Acids Research*, 1992. **20** (10): p. 2471-2483.
13. Ansorge, W., Voss, H., Wiemann, S., Schwager, C., Sproat, B., Zimmermann, J.,

- Stegemann, J., Erfle, H., Hewitt, N., and Rupp, T., *Electrophoresis*, 1992. **13**: p. 616-619.
14. Tabor, S. and Richardson, C.C., *J. Biol. Chem.*, 1990. **265** : p. 8322-8329.
 15. Ansorge, W., Zimmermann, J., Schwager, C., Stegemann, J., Erfle, H., and Voss, H., *Nucleic Acids Research*, 1990. **18** : p. 3419-3420.
 16. Swerdlow, H., Zhang, J.Z., Chen, D.Y., Harke, h.R., Grey, R., Wu, S., Dovichi, N.J., and Fuller, C., *Anal. Chem.*, 1991. **63** : p. 2835-2841.
 17. Chen, D., Harke, H.R., and Dovichi, N.J., *Nucleic Acids Research*, 1992. **20** (18): p. 4873-4880.
 18. Ruiz-Martinez, M.C., Berka, J., Belenkii, A., Foret, F., Miller, A.W., and Karger, B.L., *Anal. Chem.*, 1993. **65** : p. 2851-2858.
 19. Huang, X.C., Quesada, M.A., and Mathies, R.A., *Anal. Chem.*, 1992. **64** : p. 2149-2154.
 20. Pentoney, S.L.J., Konrad, K.D., and Kaye, W., *Electrophoresis*, 1992. **13**: p. 467-474.
 21. Nelson, M., *Nucleic Acids Research*, 1992. **20** : p. 1345-1348.
 22. Saiki, R.K., Gelfand, D.N., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A., *Science*, 1988. **239** : p. 487-491.
 23. Carothers, A.M., Urlaub, G., Mucha, J., Grunberger, D., and Chasin, L.A., *BioTechniques*, 1989. **7**(5): p. 494-499.
 24. Murray, V., *Nucleic acids research*, 1989. **17** (21): p. 8889.
 25. Hunkapiller, T., Kaiser, R.J., Koop, B.F., and Hood, L., *Science*, 1991. **254** : p. 59-67.
 26. Kieleczawa, J., Dunn, J.J., and Studier, F.W., *Science*, 1992. **258** : p. 1787-1791.
 27. Osterman, L.A., *Methods of protein and nucleic acids research*. Vol. 1. 1984, New York: Springer-Verlag.
 28. Jorgenson, J.W. and Lukacs, K.D., *Science*, 1983. **222** : p. 266-272.

29. Mikkers, F.E.P., Everaerts, F.M., and Verheggen, T.P.E.M., *J. Chromatography*, 1979. **169** : p. 11-20.
30. Ferguson, K.A., *Metabolism*, 1964. **13** : p. 985-1002.
31. Tanaka, T., *Scientific American*, 1981. **244** : p. 124-138.
32. Tietz, D., Gottlieb, M.H., Fawcett, J.S., and Chrambach, A., *Electrophoresis*, 1986. **7**: p. 217-220.
33. Rocheleau, M.-J. and Dovichi, N.J., *J. Microcol. Sep.*, 1990. **4**: p. 449-453.
34. Nishikawa, T. and Kambara, H., *Electrophoresis*, 1991. **12** : p. 623-631.
35. Smith, R.D., Olivares, J.A., Nguyen, N.T., and Udseth, H.R., *Anal. Chem.*, 1988. **60**: p. 436-441.
36. Wallingford, R.A. and Ewing, A.G., *Anal. Chem.*, 1988. **60** : p. 1973-1975.
37. Yu, M. and Dovichi, N.J., *Applied Spectroscopy*, 1989. **43** (2): p. 196-201.
38. Cheng, Y.F. and Dovichi, N.J., *Science*, 1988. **242** : p. 562-564.
39. Nguyen, D.C., Keller, R.A., Jett, J.H., and Martin, J.C., *Anal. Chem.*, 1987. **59**: p. 2158-2160.
40. Wu, S. and Dovichi, N.J., *Talanta*, 1992. **39** (2): p. 173-178.
41. Parker, C.A., *Photoluminescence of solutions*. 1968, New York: Elsevier. 411-426.
42. Wu, S. and Dovichi, N.J., *J. Chromatography*, 1989. **480** : p. 141-155.
43. Chen, D.Y., Swerdlow, H.P., Harke, H.R., Zhang, J.Z., and Dovichi, N.J., *J. Chromatography*, 1991. **559**: p. 237-246.

Chapter 2

High-Speed, High-Sensitivity DNA Sequencing By Capillary Gel Electrophoresis Based On Sheath Flow Cuvette As A Fluorescence Detector

2.1 INTRODUCTION

Small dimension separations can be traced back to the 1950's when Edström reported the separation of RNA using cellulose fibers 10 μm in diameter and 25 mm in length [1,2]. More recently, Hjerten's and Karger's groups have reported the use of gel-filled capillaries for protein separations [3,4]. Several groups have reported the use of gel-filled capillaries for separation of oligonucleotide standards with detection by ultraviolet absorbance [5-8]. An important issue is the performance of the gel-filled capillaries under high electric field strength; patents have been issued on the use of a bifunctional reagent to bind the acrylamide to the wall of the capillary to improve the gel stability [9,10].

Capillary gel electrophoresis has been applied to the separation of fluorescently labeled DNA sequencing fragments. Typical capillaries are 50- to 75- μm inner diameter (ID) and 25- to 100-cm long. Swerdlow and Gesteland reported the separation of a single dideoxynucleotide reaction mixture with a single spectral channel laser-induced fluorescence detector [11]. Drossman and co-workers in Smith's laboratory reported the separation of a single reaction mixture in gel-filled capillaries [12]. When operated at 400 V/cm the system produced sequencing rates of 1000 bases per hour after elution of the primer, a roughly 10 to 20 times higher sequencing rate than produced by conventional slab gel electrophoresis. A subsequent report from Smith's laboratory described the use of capillary gel electrophoresis for the separation of the reaction products of the four-

spectral-channel, single-lane sequencing system [13]. Resolution was a factor of 1.5 to 2 times superior to slab gel data. Karger's laboratory reported separation of DNA fragments, labeled with a single fluorescent dye, at 350 V/cm, yielding sequencing rates of 450 bases per hour; baseline resolution was obtained for fragments up to 330 nucleotides in length [14]. Gesteland's laboratory has reported the use of a charge coupled device for fluorescence detection in DNA sequencing by capillary gel electrophoresis [15]. Our laboratory has reported several designs for fluorescence detection in capillary gel electrophoresis for DNA sequencing [16-18].

High-sensitivity detection appears to be required for DNA sequencing by capillary gel electrophoresis. To ensure that the local electric field is not perturbed by the ionic strength of analyte, it is necessary that the ionic strength of the sample be less than 1% of the ionic strength of the buffer. Although the buffers used in gel electrophoresis have concentration in the high-millimolar range, they are poorly ionized and the total ion concentration of the buffer is about 10 mM. Therefore, the total ionic strength of the sample in the electrophoresis gel must be less than 100 μ M, to minimize electric field artifacts. At least half the analyte molecules are unlabeled template. Because the template, which is about 7500 bases long and has a negative charge per base, carries a high charge, most of the ionic strength of the sample is associated with template. If one assumes that the average analyte fragment is 300 bases in length, only 4% of the ionic strength of the sample is associated with analyte; the analyte ionic strength is 4 μ M. However, since each analyte molecule has, on the average, 300 bases, the concentration of fluorescently labeled analyte is about 1×10^{-8} M. Clearly, very sensitive detection is required for DNA sequencing.

The sample must occupy a volume less than 10^{-9} l in the capillary gel to avoid excessive band broadening; an average of 10 attomole (1 attomole = 1 amol = 10^{-18} mol = 600,000 molecules) of DNA will be present in each band. Although additional material can be loaded onto the column through use of stacking injection conditions, only a minute

amount of material is available for sequencing. To obtain a decent signal/noise ratio and to account for those fragments that are incompletely labeled, it is necessary to use a system with zeptomole (1 zeptomole = 1 zmol = 10^{-21} mol = 600 molecules) detection limits in DNA sequencing by capillary gel electrophoresis. On-column fluorescence detectors, for which the capillary serves as the detection chamber, produce detection limits of about 200 zmol of fluorescently labeled product [11,12,14,16]. These on-column detectors are dominated by a large background signal generated by light scatter from the capillary walls, along with scatter from scratches and pits in the wall. Scatter takes the form of a fan of light in the plane perpendicular to the capillary axis. Although detection of the scattered light is minimized by tilting the capillary relative to the collection optic, use of a high numerical aperture objective to increase collection efficiency of fluorescence also enhances sensitivity to scattered light. Refraction and reflection may be minimized by use of a tightly focused laser beam so that only a small fraction of the capillary cross section is illuminated; although background is reduced, only a small fraction of analyte molecules eluting from the electrophoresis column is illuminated, decreasing sensitivity. Furthermore, the high irradiance produced by the focused beam leads to enhanced photobleaching of the analyte, an important problem with the photolabile dyes used in the sequencing reactions. To eliminate light scatter associated with the capillary windows, a post-column fluorescence detector can be used. The detector, based on a sheath-flow cuvette, can produce excellent detection performance at modest cost and effort.

2.2. INSTRUMENTAL

2.2.1. DESIGN CONSIDERATION

A sheath flow cuvette was used as the detection chamber for capillary

electrophoresis (Fig. 2.1). This chamber is typically 200- μ m square in cross section, 20-mm long, and has 2-mm thick windows. The quartz windows have good optical quality and generate very little light scatter. The 190- μ m outer diameter (OD) capillary is placed within the square chamber. As analytes elute from the bottom of the capillary, they are swept downstream by a sheath flow, which is usually composed of the same buffer as is used for the separation. Sheath flow rate is low, about 0.15 ml/h. The sheath stream is generated with a high-precision syringe pump or by gravity-driven siphon flow.

Disruption of the flow pattern within the cuvette occurs when drops dislodge from the bottom of the cuvette. A piece of plastic tubing can be used to direct the waste flow from the cuvette to a receiving vessel, eliminating the formation of droplets. In capillary zone electrophoresis, it is important that the sample and waste reservoirs are at the same height; otherwise, a siphon can form between the two reservoirs. The siphon will lead to bulk flow that both biases the injection volume and introduces a parabolic flow profile in the separation capillary, distorting the separation. The high resistance to flow found in gel capillary eliminates hydrodynamic flow in the separation capillary when the receiving and injection reservoirs are not at the same height.

At the very low flow rates employed in the sheath-flow cuvette, flow is laminar. Analyte forms a narrow stream in the center of the flow chamber, far from the cuvette walls. By restricting the field of view of the photodetector to the sample stream, light scatter from the interface of the cuvette window and sample stream is not detected, greatly reducing the background signal. In capillary electrophoresis, detection limits of single molecules have been achieved in this laboratory [19].

The DuPont-sequencing technique uses succinylfluorescein dyes to label the four dideoxynucleotides (see Fig. 1. 3 in Chapter 1). These dyes have closely spaced absorbance and emission bands. As a result, a single argon ion laser wavelength, 488 nm, can be used to excite fluorescence from all dyes. Emission is distributed between two spectral channels centered at 510 and 540 nm. The ratio of the fluorescence intensity in

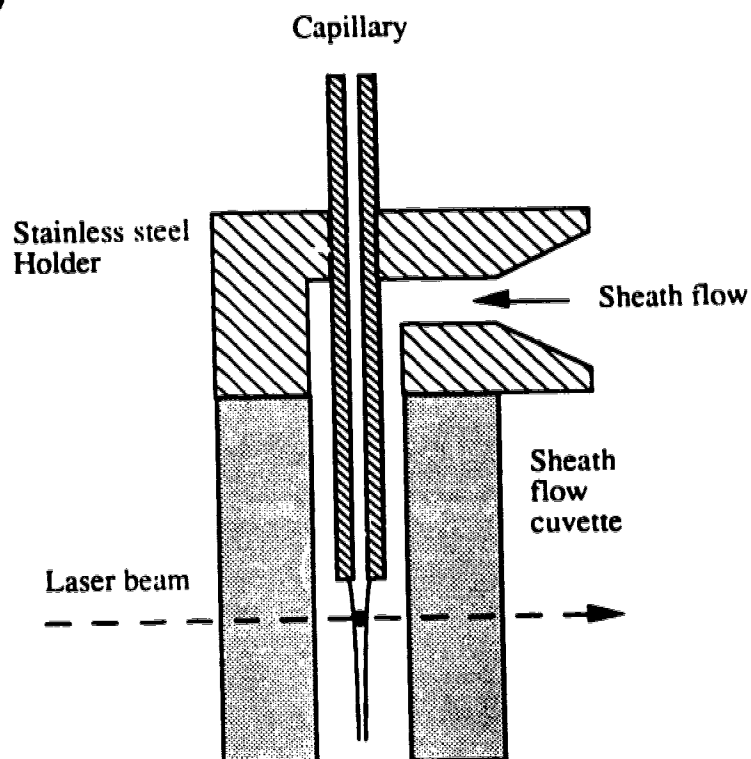
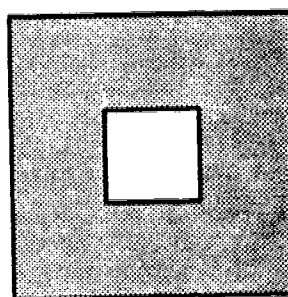
a. Side view**b. Top view**

Fig. 2.1. A schematic of the sheath flow cuvette as a post-column fluorescence detector. The sheath flow draws the sample stream down and the laser beam excites fluorescence from the sample stream below the exit of the capillary.

the two spectral channels is used to identify the terminating dideoxynucleotides. Because of photobleaching, modest laser power is used in these experiments. The longest-wavelength dyes used in the DuPont-sequencing protocol are particularly sensitive to photobleaching; typically 5-mW laser power is used to excite fluorescence from all dyes.

It would be ideal to use two sides of the sheath flow cuvette for the fluorescence detection, one for the spectral channel centered at 510 nm and the other for the spectral channel centered at 540 nm. There is a technical difficulty in using the sheath flow cuvette in this way. Only one side of the sheath flow cuvette is used for fluorescence collection. The opposite side to the collection optics is occupied by an auxiliary microscope, which is used to assist the alignment of the optical system.

Fluorescence is collected at right angles to the laser beam with a high-efficiency microscope objective and imaged onto a pinhole to restrict the view of the illuminated sample stream. A single band-pass filter is used to block scattered laser light, whereas a dichroic filter is used to split the fluorescence into two spectral channels, followed by detection with two photomultiplier tubes (PMTs).

2.2.2. INSTRUMENT

The instrument was built on an optical breadboard, 3 feet \times 4 feet (Melles Griot, CA, USA), as shown in Fig. 2.2. The sheath flow cuvette, 200 μm \times 200 μm in cross-section and 20 mm in length (See Fig. 2.1. Precision Cells, New York, USA), was installed in a home-made sheath flow cuvette holder, which was mounted on three-axes translation stages. The sheath flow was supplied with a syringe pump, Model 314 (Isco, Lincoln, NE, USA). The argon ion laser (Uniphase, CA, USA), operated in the light regulated mode, was used to produce a beam at 488 nm. Usually laser power of less than 10 mW was generated for the experiment. This beam was reflected by a flat surface-coated mirror (Newport, CA, USA) and focused by a 30.8 mm focal length, 4 \times

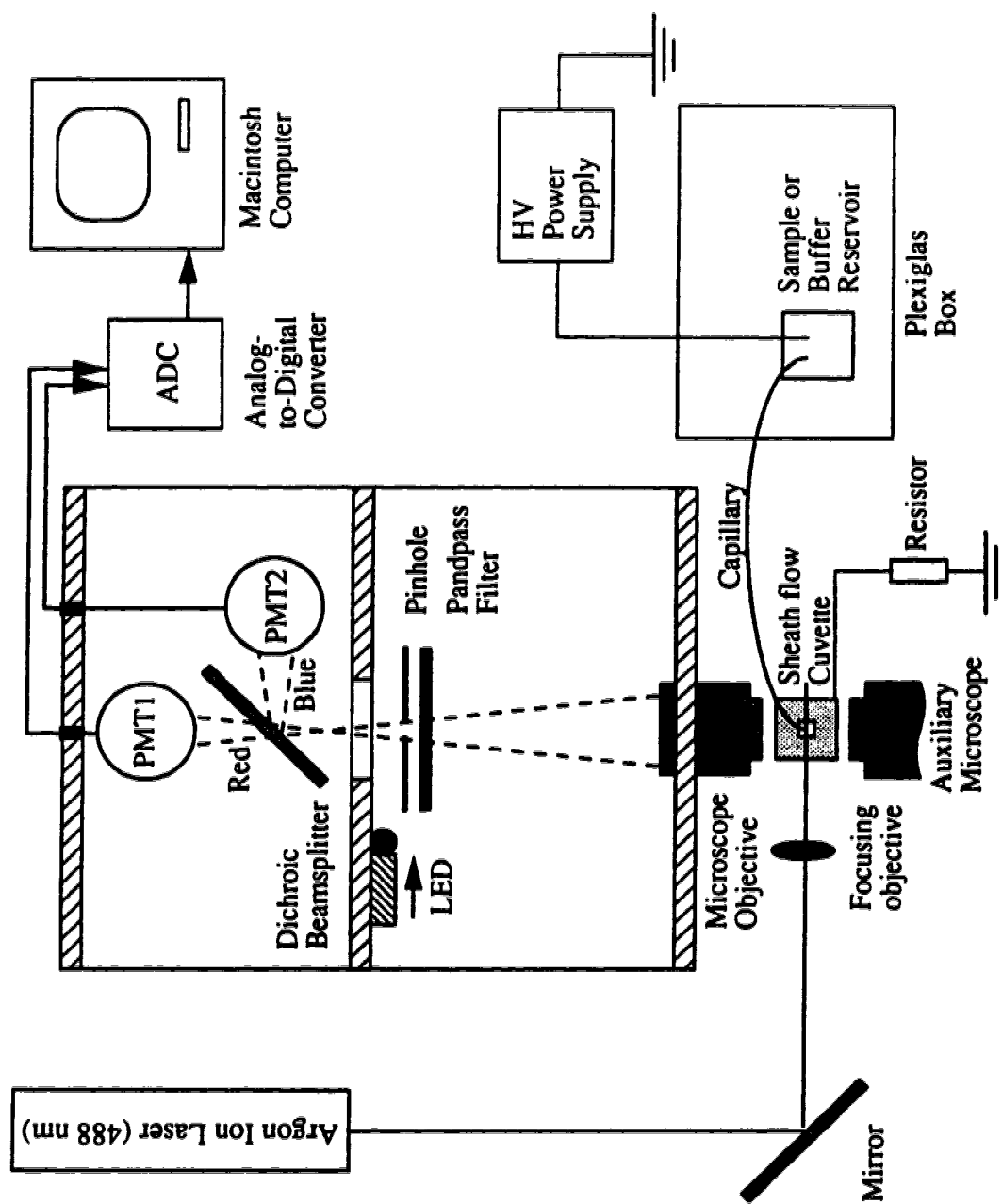


Fig. 2.2. Sheath flow cuvette and two-spectral channel laser-induced fluorescence detection system.

microscope objective (Melles Griot, CA, USA), which was also mounted on another three-axes translation stages, onto the sheath flow cuvette to excite fluorescence from sample streams. Three aluminum plates, 8 in. \times 8 in. \times 1/4 in. were used to build two compartments with a 1 in. diameter hole drilled on both the front and middle plates; the distance between front plate to the middle plate was 170 mm and the distance from the middle plate to the back plate was 100 mm. The fluorescence-collecting microscope objective, 0.65 NA/32 \times (Leitz, Calgary, Canada), was mounted on the front plate through the 1-in. hole; a green light-emitting diode (LED, which was used to check the position of pinhole), driven by a solenoid, was mounted on the middle plate in front of the 1-inch hole. Photomultiplier tube (PMT) power supply connectors and PMT signal connectors were mounted on the back plate. In the first compartment was an interference bandpass filter, 530 nm center and 45 nm passband (Omega, VT, USA), and a 600- μ m pinhole, about 150 mm away from the microscope objective. In the second compartment were the dichroic beamsplitter, 525DCLP (Omega, VT, USA) and two PMTs (Hamamtsu, CA, USA). One PMT was behind the beamsplitter and the other PMT was on the right side of the beamsplitter. Aluminum plates and all the optical parts were fixed on top of the optical breadboard to form a simple, solid design. Except the collection objective, all these optics were covered with 1-mm thick black cardboard, which was supported by the aluminum plates, to form a dark box.

The 488 nm line from an argon ion laser is very close to 490 nm, the maximum absorbance of fluorescein, and the bandpass filter, 530DF45, is chosen to match the fluorescence emission spectrum, which has its maximum emission at 518 nm. The wavelength profiles of the laser line and bandpass filter also met the requirement for excitation and detection of the four succinylfluorescein dyes used in Dupont's protocol [20].

It is the spectra of the bandpass filter and the dichroic beamsplitter that define the two spectral channels (Fig. 2.3), the design of which is specific for the DuPont

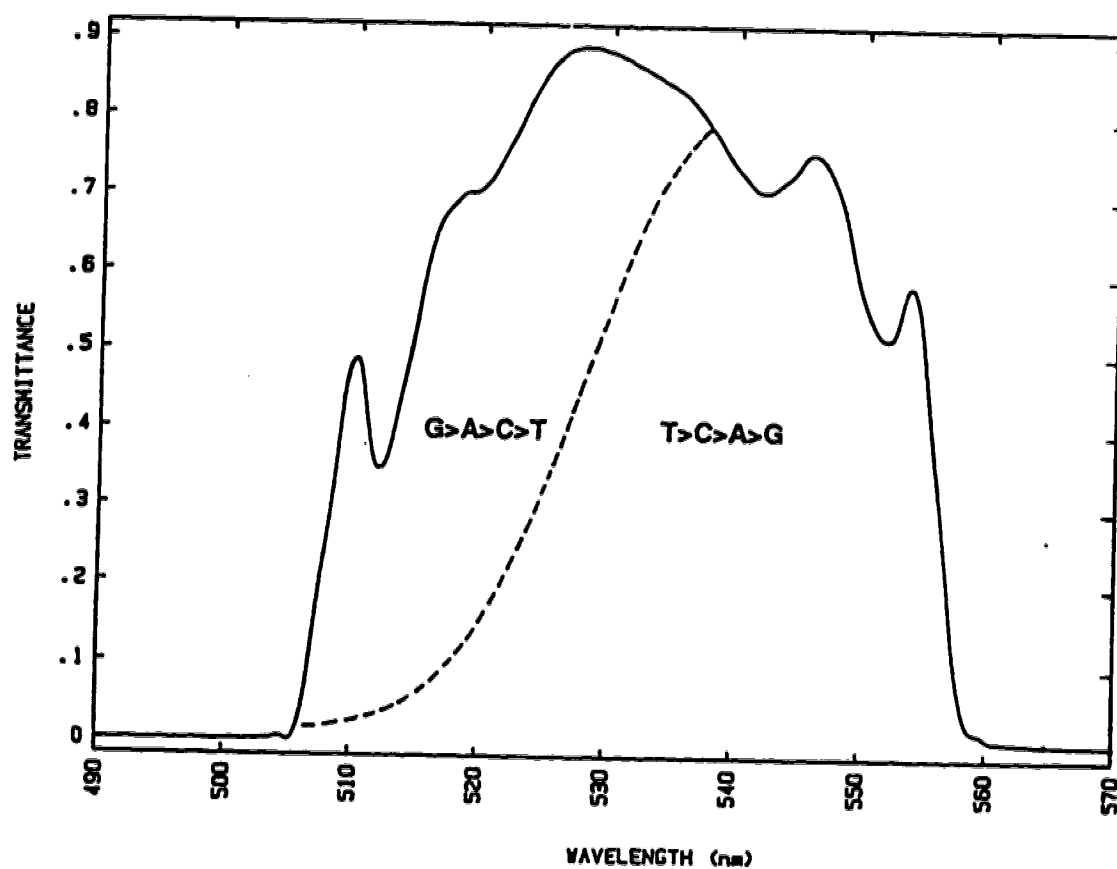


Fig. 2.3. The spectra of the bandpass filter, 530DF45 (solid line), and the dichroic beamsplitter, 525DCLP (dashed line). The two spectral channels are outlined, with G > A > C > T in the blue channel and T > C > A > G in the red channel.

sequencing protocol. The dichroic beamsplitter (dashed lines) divided the passband into two parts: one is centered at 510 nm to form a blue channel, while the other is centered at 540 to form a red channel. Normalized fluorescence intensity should be $G-505 > A-512 > C-519 > T-526$ in the blue channel; whereas normalized fluorescence intensity should be $T-526 > C-519 > A-512 > G-505$ in the red channel. The ratio of the fluorescence intensities in the blue channel to the fluorescence intensity in the red channel should be in the order of $G-505 > A-512 > C-519 > T-526$.

An auxiliary microscope was placed on the side opposite the collection optics to inspect the detection system alignment. To protect the operator, there was a colored glass filter, OG 515 (Melles Griot, CA, USA), fitted in this microscope. In this design, the collecting optic was fixed while the optical alignment was achieved by adjusting the positions of the sheath flow cuvette as well as the focused laser beam. The alignment was performed in two steps. First, an alignment solution of fluorescent dye exiting the end of the capillary is excited by the laser beam. Then the LED is used to illuminate the pinhole from behind. Through the auxiliary microscope, the position of the fluorescent spot is adjusted so that this spot is superimposed with the illuminated pinhole. Second, the LED is removed from the behind of the pinhole, and the relative position of the laser beam and the sheath flow cuvette is adjusted until a maximum fluorescence signal is reached.

The two PMT's were powered by a single power supply (Model 204-10, Pacific Instruments, Concord, CA, USA). The PMT outputs were conditioned with a low-pass electronic filter, 0.5-s (500 k Ω resistor and 1 μ F capacitor) time constant, and then digitized and stored by a Macintosh Plus computer.

One end of the capillary was inserted into the sheath flow cuvette. The other end of the capillary was inserted in a buffer vial that was confined in a Plexiglas box equipped with a safety interlock. High voltage was provided by a high voltage power supply (0 - 30 kV, Spellman, Plainview, NY, USA) through a platinum electrode that was also inserted into the buffer vial. The electric circuit was completed by connecting the

sheath flow cuvette holder to ground through a $\sim 100\text{ k}\Omega$ resistor, which served to monitor the current change during the process of electrophoresis.

2.3. EXPERIMENTAL

2.3.1. MATERIALS

Fluorescein was a high purity standard from Molecular Probes (OR, USA). Acrylamide, N, N'-methylene bisacrylamide (Bis) and N,N,N',N'-tetramethylethylenediamine (TEMED) were electrophoresis-purity reagents from Bio-Rad (Mississauga, Canada). Tris base and urea were ultrapure reagents (ICN, Montreal, Canada). Borax, boric acid and EDTA were analytical-reagent grade (BDH, Ontario, Canada). Ammonium persulfate was ultrapure electrophoresis grade (Boehringer Mannheim, Laval, Canada). γ -methacryloxypropyltrimethoxysilane was from Sigma (St. Louis, MO, USA). Capillaries, 50- μm ID and 190- μm OD from Polymicro Technologies (Phoenix, AZ, USA), were used as received with no pretreatment.

The $1 \times$ TBE (Tris-boric acid-EDTA) buffer, pH8.2, is the ten-fold dilution of $10 \times$ TBE that was made by dissolving 5.4 g Tris, 2.75 g boric acid and 0.48 g EDTA in water to a total volume of 50 ml. Fluorescein was first dissolved in ethanol to make the concentration in the range of 1-mM. This solution was diluted to a desired concentration in borate buffer, 10 mM, pH9.2. The diluted fluorescein solution was used within 24 hours.

2.3.2. CAPILLARY GEL PREPARATION

The capillary was first treated with 1% γ -methacryloxypropyltrimethoxysilane solution (95% ethanol-5% water, pH 4.8), either by filling the whole capillary with a

syringe or simply by dipping one end (usually the detection end) of the capillary into the solution for 1 minute. A 4%T, 5%C acrylamide solution was prepared by adding 10.5 g urea, 2.5 ml 40% premixed stock acrylamide-Bis solution (19:1) and 2.5 ml 10 × TBE (890 mM Tris-borate, 20 mM EDTA) into water to make a total volume of 25 ml, pH 8.3. This solution was filtered with a 0.22- μ m filter unit (Millipore, Millipore Indústria e Comércio Ltda, Brazil). Five ml of this filtered solution was degassed with a water aspirator for 20 minutes. Two μ l of TEMED and 20 μ l of 10% ammonium persulfate water solution were added to the acrylamide solution to initiate polymerization. The acrylamide solution was immediately injected into the capillary with a syringe. The solution was introduced until about 10 to 20 tiny drops of acrylamide solution dislodged from the other end of the capillary onto a piece of Kimwipe. Polymerization was allowed to proceed overnight. The gel-filled capillaries were checked under microscope and the bubble-free capillaries were chosen for DNA fragment separation. About 90% of the capillaries were bubble-free.

2.3.3. DNA SEQUENCING SAMPLE

The sample was prepared from a DuPont Genesis 2000 protocol, using 3 μ g of M13mp18 single-stranded template, which was annealed to 15 ng of -40 17 mer M13 primer, and 1 μ l of Sequenase (U.S. Biochemicals) in a standard reaction mix (3 μ l 75- μ M deoxy solution, 2.5 μ l 0.1-M dithiothreitol, 1 μ l Sequenase dideoxy mix). The reaction products were ethanol precipitated, washed, and resuspended in 3 μ l of 49:1 mixture of formamide:0.5 M EDTA at pH 8.0. The sample was heated at 95 °C for 2 minutes prior to being loaded into the gel-filled capillary. The sample was electrokinetically loaded onto the gel-filled capillaries, typically at 100 V/cm for 30 seconds.

2.4. RESULTS AND DISCUSSION

2.4.1. LIMITS OF DETECTION (LOD)

The detection limit of the instrument was characterized in the free zone electrophoresis format. Fluorescein was diluted in 10 mM borate buffer to a final concentration of 3.2×10^{-11} M. This diluted fluorescein was injected at 500 V for 10 seconds and the electrophoresis was driven at 400 V/cm for 4 minutes over a 50- μ m ID, 38-cm long capillary. Four measurements were made and the results are shown in Table 2.1. A relative standard deviation of 5-6% was obtained for the 4 measurements. Fig. 2.4 presents one of the measurements in the two spectral channels. Blank measurements showed no peaks.

Table 2.1. Two-spectral-channel fluorescence detection of a 3.2×10^{-11} M solution of fluorescein molecules

#Run	Peak height (Arbitrary unit)	
	Red Channel	Blue Channel
1	128	100
2	126	96
3	117	88
4	114	95
\bar{x}	121	95
σ	7	5
$\sigma/\bar{x}(\%)$	6	5

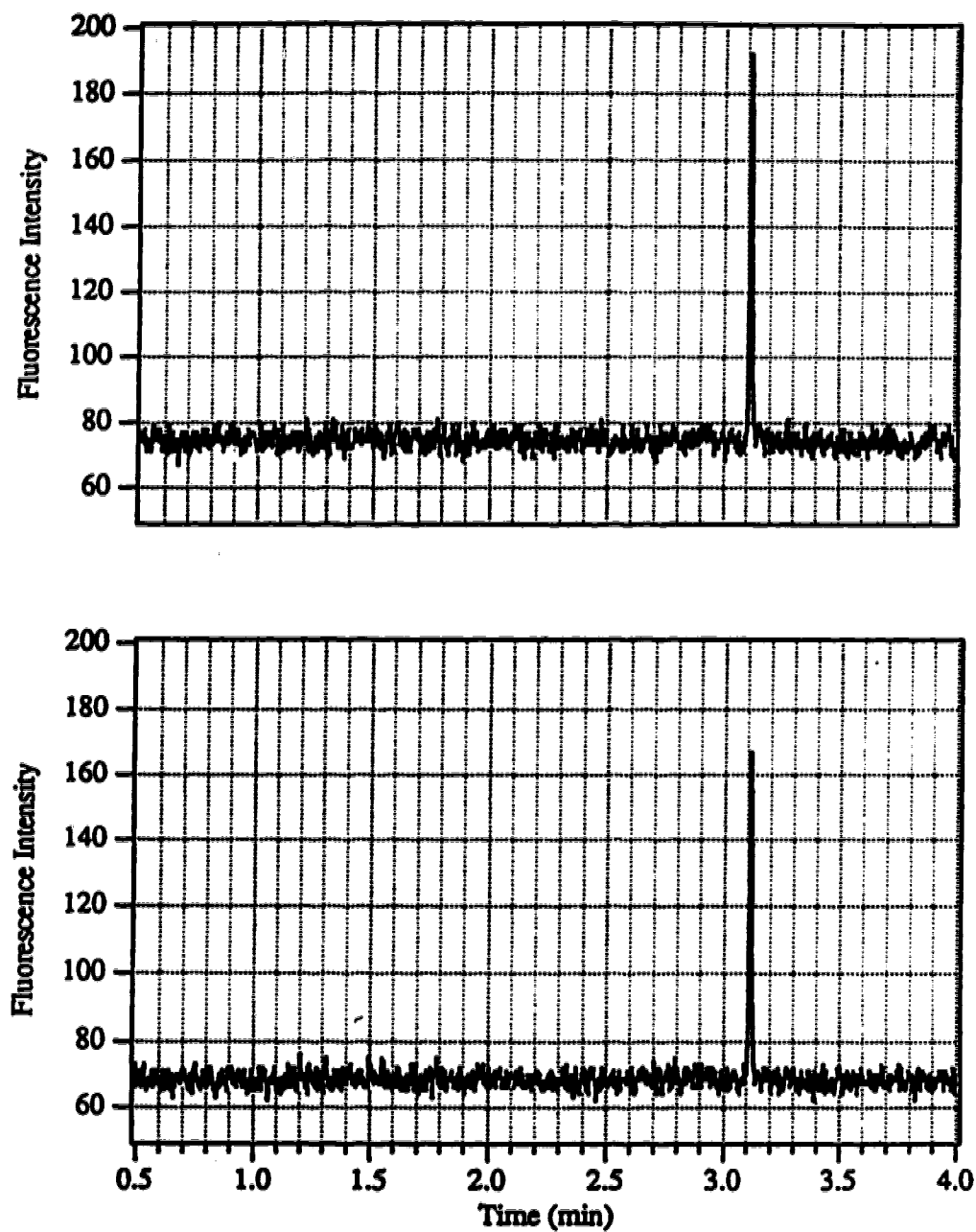


Fig. 2.4. Detection limit measurements in the two spectral channels. The top is the fluorescence signal in the blue channel and the bottom is the fluorescence signal in the red channel. Peaks are from injection of 1.4 nl of a 3.2×10^{-11} M solution of fluorescein.

Knoll proposed a one-point method for estimating LOD of a chromatographic procedure [21]. This method assumes that chromatographic peaks are Gaussian. It defines LOD as the analyte concentration that produces a chromatographic peak having a height equal to three times the standard deviation of the baseline noise (3σ). The concentration detection limit, C_{LOD} , for a Gaussian peak is given by

$$C_{\text{LOD}} = K_{\text{LOD}} h_n C_s / h_s \quad (2-1)$$

where C_s is the concentration of the calibration sample, h_s is the measured peak height, h_n is the height of the largest noise fluctuation in a pre-selected interval, and K_{LOD} is a constant and its value depends on the length of the pre-selected interval, which is the multiple of the full width at half height of the analyte peak (Table 2.2).

According to this method, the concentration detection limit, C_{LOD} , is calculated to be 0.9×10^{-12} M for the red channel and 1.1×10^{-12} M for the blue channel. The injected volume, V_{inj} , can be estimated by the following equation [22]:

$$V_{\text{inj}} = \frac{V_i \times t_i}{V_{\text{el}} \times t_m} V_c \quad (2-2)$$

where V_i is the injection voltage, V_{el} is the electrophoretic voltage, t_i is the injection time and t_m is the migration time of fluorescein at the given condition. The injection volume is calculated to be 1.4×10^{-9} l or 1.4 nl. The mass detection limits were 1.3 and 1.5 zeptomole, or 760 and 930 fluorescein molecules injected onto the capillary for the red

Table 2.2. Some values K_{LOD} of for different interval lengths

Peak width multiple	10	20	50	100
K_{LOD}	1.9718	1.4309	0.9194	0.6536

channel and the blue channel, respectively. This result is comparable to that reported previously, where the best LOD was 600 fluorescein-labeled arginine molecules with detection in only a single spectral channel [23].

2.4.2. DNA SEQUENCE DETERMINATION

Electrophoretic separations were carried out in a capillary filled with 4%T, 5%C polyacrylamide gel. A very low sheath flow rate, 0.16 ml/h, was employed to transport analyte from the capillary to the illuminated region. During the transit time, analyte can diffuse an appreciable amount, decreasing the effective concentration of analyte in the detector. However, higher-molecular-mass DNA fragments undergo less diffusion and are more concentrated in the illumination volume. As a result, the detection limit for the system improves for larger DNA fragments, which are typically produced in sequencing reactions at lower concentration compared with early-eluting fragments. Fig. 2.5 shows a separation of the fragments from an M13mp18 template at 287 V/cm. The first large peak, eluting at 19.1 minutes is associated with unincorporated dideoxynucleotides. The peaks from 19.4 to 20.5 minutes are difficult to interpret and are associated with primer. The large peak at 21.4 minutes is a severe compression associated with a GC-rich portion of the sequence. The remainder of the electropherogram is consistent with the known sequence of M13mp18 to about 250 bases. The signal amplitude for longer fragments is quite low, and the data are difficult to interpret. The decrease in signal for longer fragments is associated with poor incorporation of the dideoxynucleotides in the reaction mixture. The incorporation of the fluorescently labeled dideoxynucleotides is not so efficient as the incorporation of their unlabeled counterparts. This condition leads to termination of the reaction at relatively short fragments, so that there is very poor signal amplitude for the longer fragments. Alternative sequencing protocols can be employed that allow much longer sequencing runs to be generated. The data presented here are

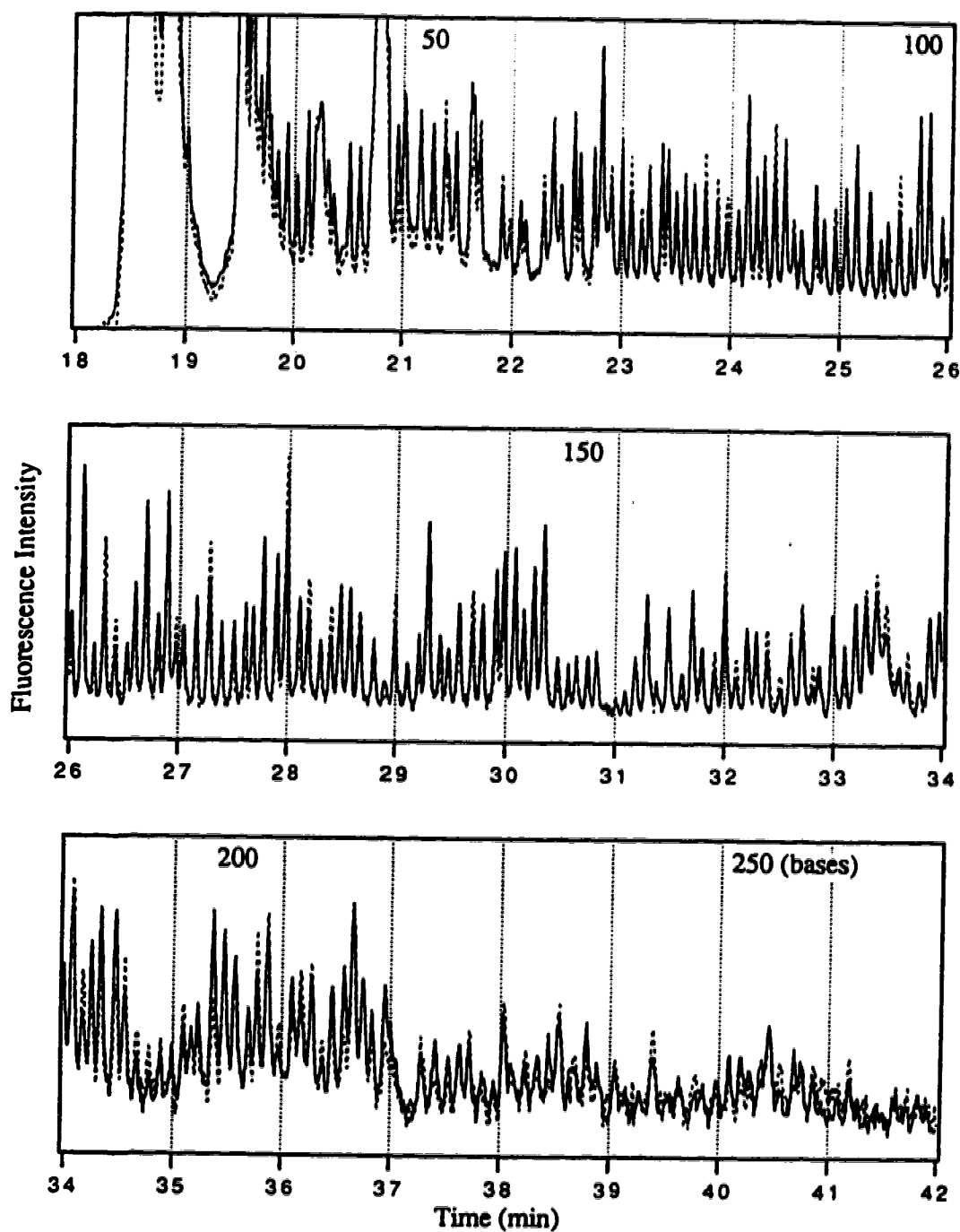


Fig. 2.5. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection. Solid line is the signal from the red channel and dashed line is the signal from the blue channel. The electrophoresis was conducted at 287 V/cm.

simply illustrative of the two-spectral-channel detector employed with this particular sequencing protocol and the behavior of the gel in the separation under high electric field strength.

The same gel was used for another sequencing separation at a much higher electric field strength, 461 V/cm (Fig. 2.6). The electropherogram is similar to Fig 2.5, but fragments 250 bases long eluted in 25 minutes, compared with 40 minutes at 287 V/cm. Elution time is nearly proportional to the applied electric field.

The peak spacing in Fig. 2.5 and Fig. 2.6 is quite uniform; Fig. 2.7 is a plot of migration time versus fragment length, which yields straight lines ($r = 1.000$) for both runs. Sequencing rate is defined as the reciprocal of the slope: at 285 V/cm, the sequencing rate is $1/0.0973 = 10.3$ bases per minute, or 620 bases per hour, for fragments 50 to 250 bases in length. At 465 V/cm, the rate is $1/0.0490 = 20.4$ bases per minute, or 1220 bases per hour, for fragments 50 to 250 bases in length; this rate is among the highest achieved in DNA sequencing in crosslinked polyacrylamide gel.

It seems that Equation 1-2, which indicates that migration rate is proportional to the electric field strength, only holds for capillary gel electrophoresis to some extent. If the sequencing rate is 620 bases per hour at 287 V/cm, it should be 996 instead of 1220 bases per hour at 461 V/cm. The difference is 18 %, which implies that temperature in the gel operated at 461 V/cm is about 8 °C higher than that operated at 287 V/cm, because mobility will increase by 2.3% with one degree increase in temperature [24].

Another gel-filled capillary was made from the same acrylamide solution for a subsequent run at 465 V/cm. The result is shown in Fig. 2.8. The DNA sequencing rate is 17 bases per minute or 1020 bases per hour. The difference is 1.4% (or a temperature rise of 0.6 °C) compared to that obtained at 287 V/cm. This result seems in agreement with Equation 1-2.

For newly prepared gel, a diluted fluorescence solution is required to align the optical system; it takes about an hour at 200 V/cm. During the process of optical

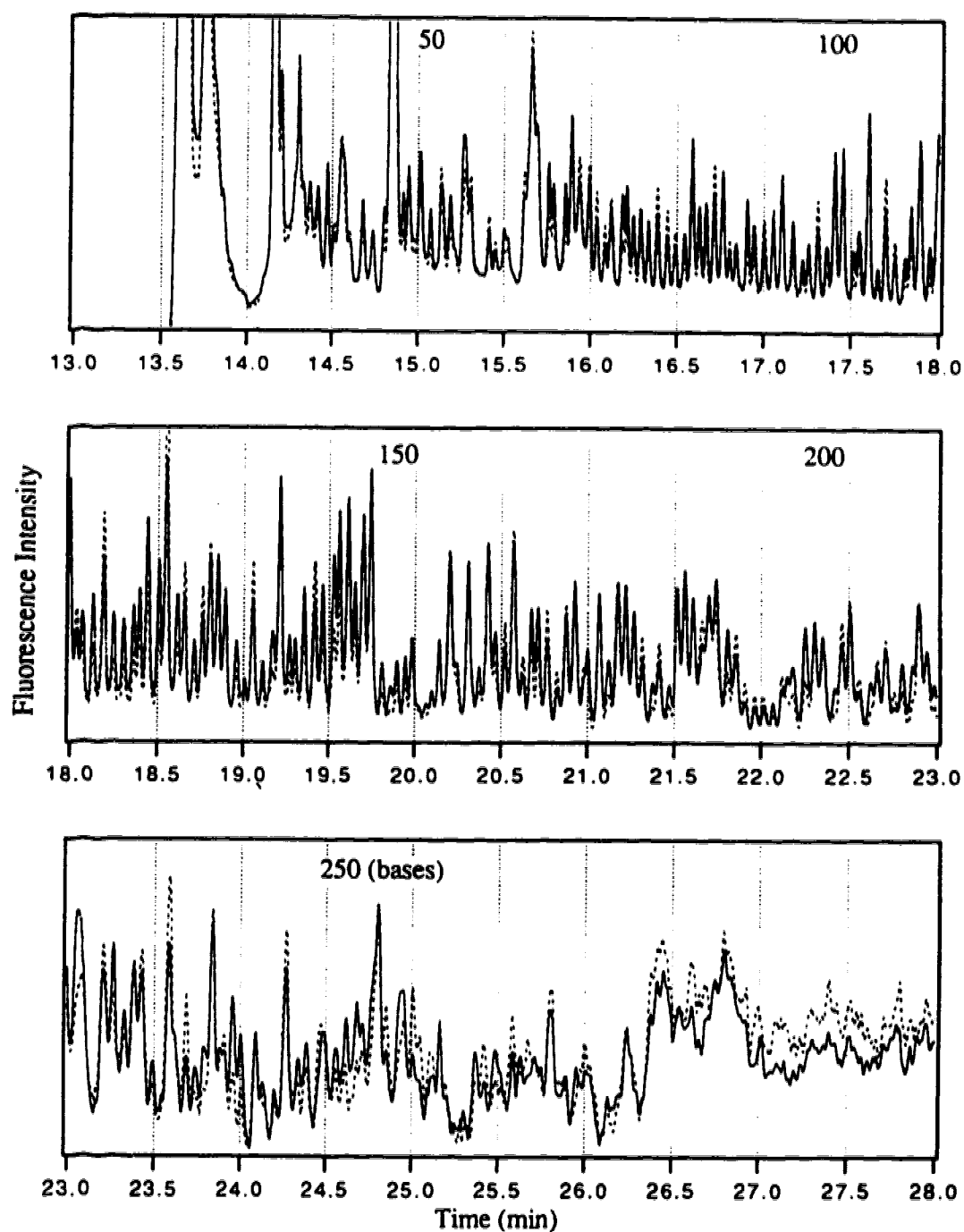


Fig. 2.6. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection. Solid line is the signal from the red channel and dashed line is the signal from the blue channel. Electrophoresis was conducted at 461 V/cm.

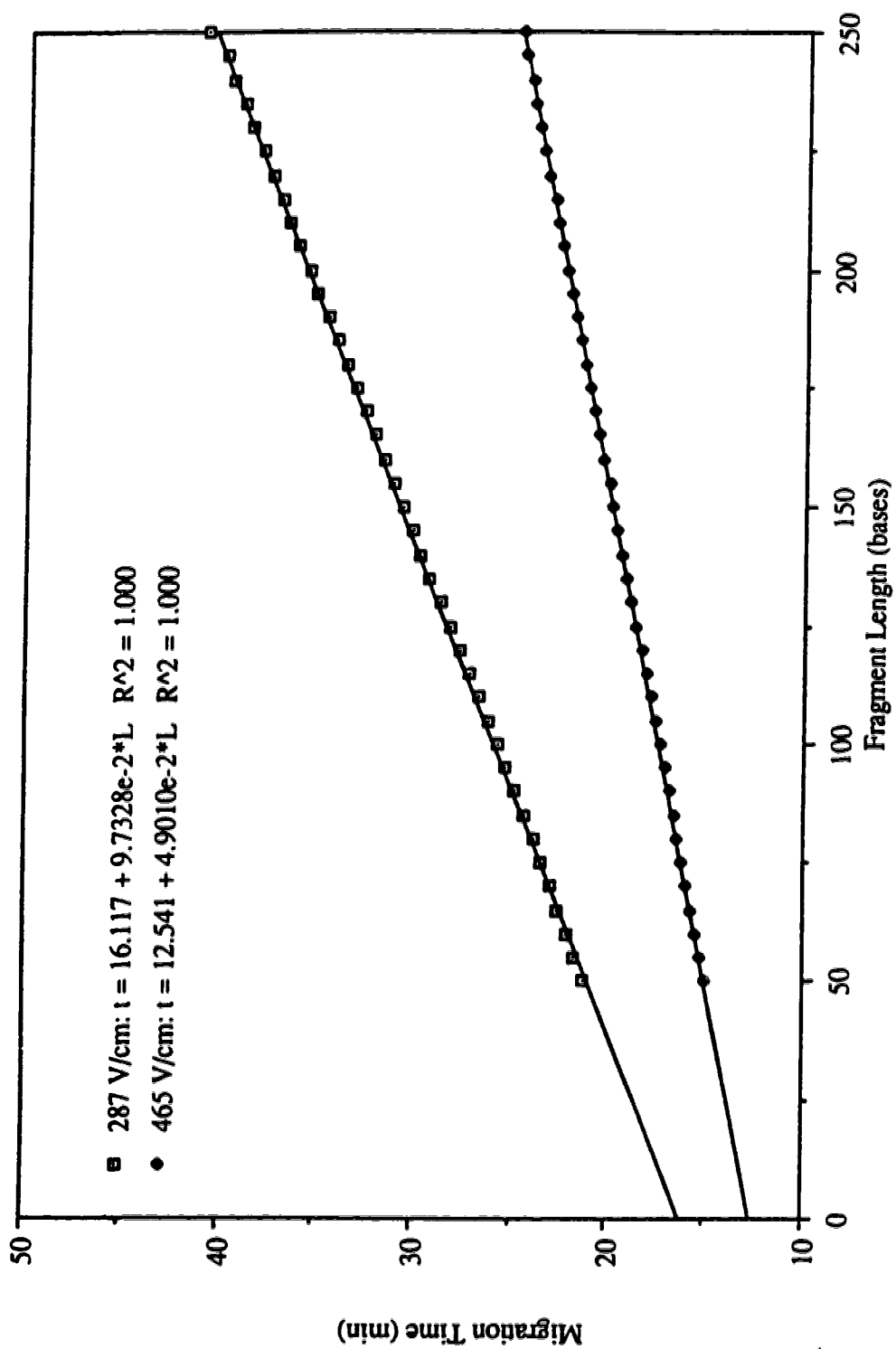


Fig. 2.7. Migration time as a function of fragment length for single-stranded DNA fragments separated in a 4% T, 5% T gel at different electric field strengths. Straight line is the linear least square fit to the data.

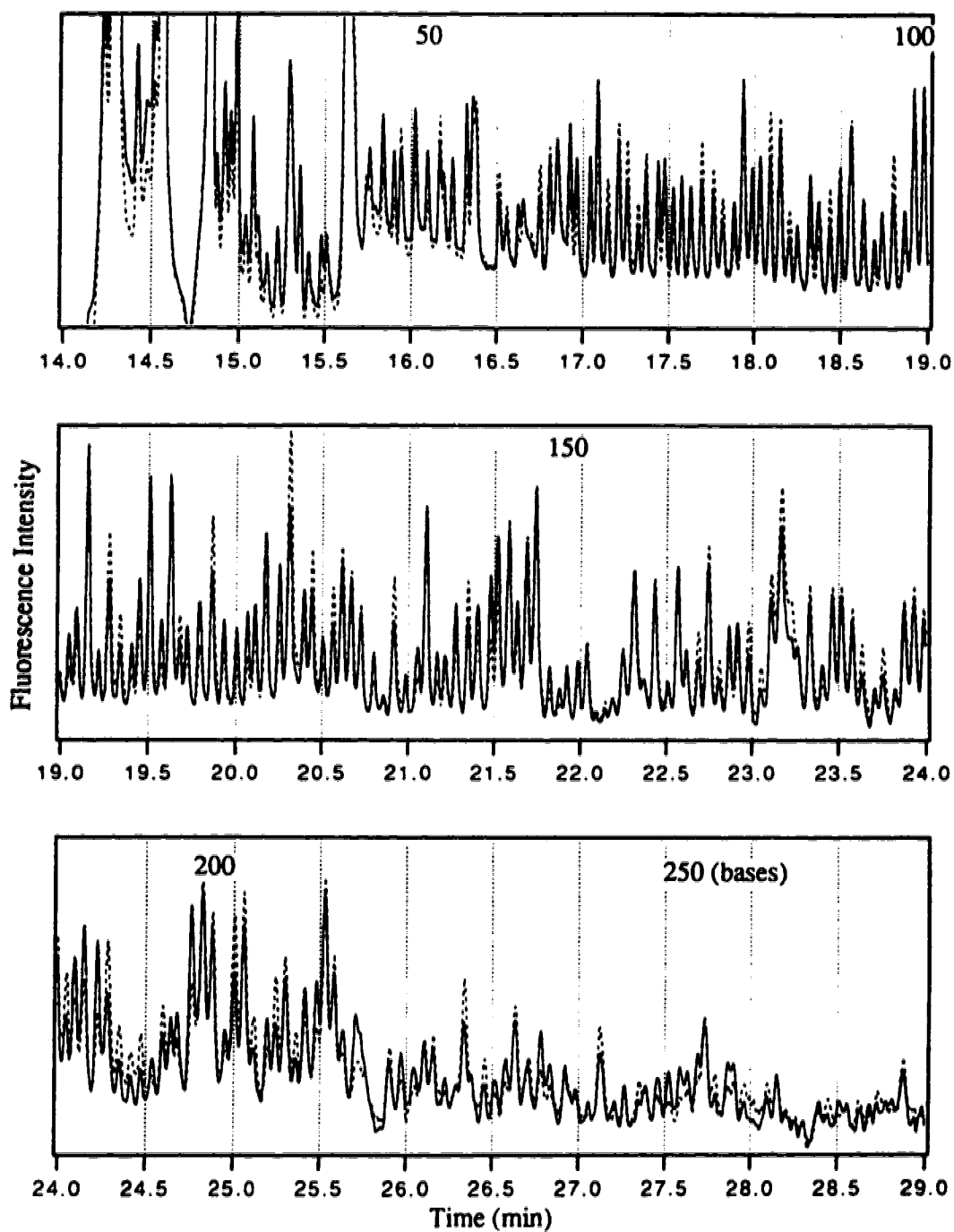


Fig. 2.8. DNA sequencing by capillary gel electrophoresis and two-spectral channel fluorescence detection. The solid line is the signal from the red channel and the dashed line is the signal from the blue line. Electrophoresis was conducted at 465 V/cm.

alignment, the electrophoretic current decreases significantly, presumably due to the induced field build-up at the gel-buffer interface [25], which is opposite in direction and substantially reduces the effective electric field strength to drive the electrophoresis. This induced electric field can be diminished with recovery time after a sufficient recovery time at zero field. As a consequence, an increase in current (Fig. 2.9) is observed at the same electric field strength. DNA fragment separation will experience a reduced effective electric field strength if the sample was loaded immediately after the optical alignment. Both Fig. 2.5 and Fig. 2.8 were generated in this manner and, as a result, they are comparable. However, in generating the data for Fig. 2.6, the injection was the second performed on the gel, and there was a 12-hour recovery period prior to the injection, leading to a relatively higher effective electric field strength, and hence higher sequencing speed, for the separation.

According to Equation 1-7, resolution should scale with square root of the applied voltage, or $R_s \propto \sqrt{V}$. This is not true as we compare the data generated at 287 and 465 V/cm from bases 84 to 87. Resolution of 1.5 is similar at higher voltage to the resolution of 1.7 at low voltage. The temperature might have some effect on resolution of the separation run at very high voltage, but the major effect is from the quality of the gel. Separation of DNA sequencing fragments at over 460 V/cm eventually destroyed gel. Larger bubbles were formed at the injection end and smaller bubbles were found along the entire length of the capillary. Those small bubbles might be the source of the band-broadening.

Gel stability is a big issue for DNA sequencing under high electric field. In the process of gel preparation, the capillary was silanized to increase the gel stability; the acrylamide solution was carefully filtered and degassed before filling in the capillary; the gel-filled capillary was checked under a microscope to make sure the gel was bubble-free. However, the chance of bubble formation exists during the process of gel electrophoresis, particularly at the sample injection end [25]. The higher the electric field strength that is

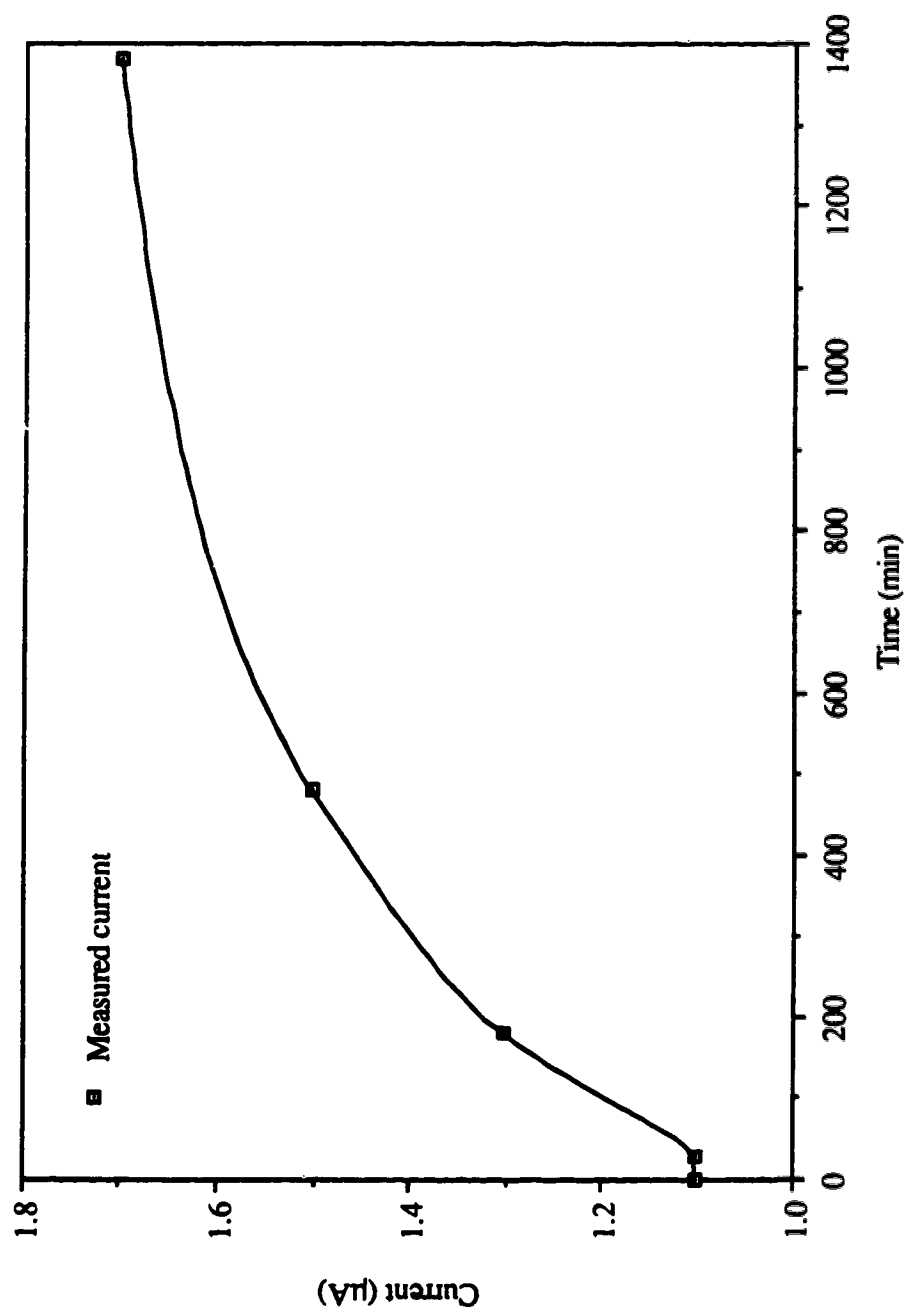


Fig. 2.9. A current increase as a function of recovery time. After an initial electrophoresis, the capillary gel was regenerated as reflected by a slow increase in current. For each data point, the current was measured at 200 V/cm for 20 minutes over a 20-cm capillary. Smooth curve is drawn as a visual aid.

applied, the greater the chance of bubble formation. Bubble formation usually happens immediately after sample injection, presumably caused by the presence of large DNA fragments such as the templates [16,25]. A small bubble can be a catastrophic site that leads to formation of large bubbles and may eventually destroy the gel. In the case of bubble formation, there is a sudden decrease in current that is flowing through the capillary. When this happens, 1- to 2-mm of the capillary must be trimmed from the injection end, the voltage is adjusted to the new capillary length, and the electrophoretic separation is resumed. Current is normally stable afterward, even though it slowly decreases during the whole separation process [25].

To optimize the gel performance, separation potential was increased in a stepwise fashion at the first several minutes until the desired potential was reached. The above two DNA fragment separations were carried out in this fashion. At the time of sample injection at 100 V/cm there was a slight increase in current. After sample injection, the sample was replaced with a fresh buffer solution, or $1 \times$ TBE. Bubble formation and current disruptions often happened immediately after the high voltage was applied, and 1-mm of capillary was trimmed after 1-minute of run. The electric field was adjusted and separation was resumed at 100 V/cm followed by increasing the potential in a stepwise fashion to 16 kV (465 V/cm, Fig. 2.10). The stability of the current at every step was used to decide whether the capillary should be trimmed or the potential should be further increased. This stepwise voltage increase tended to generate high quality separations and increase the opportunity for separation of DNA sequencing fragments at very high electric field strength.

In the region from 60 to 100 nucleotides, the peaks are nearly baseline resolved (Fig. 2.11). Resolution ranges from 1 to 2 and typically is 1.5. In this sequencing protocol, the relative peak amplitude in the two spectral channels is used to identify the terminal nucleotide. In Fig. 2.8b, the solid line corresponds to a fluorescent signal detected in the red channel centered at 540 nm, whereas the dashed line corresponds to a

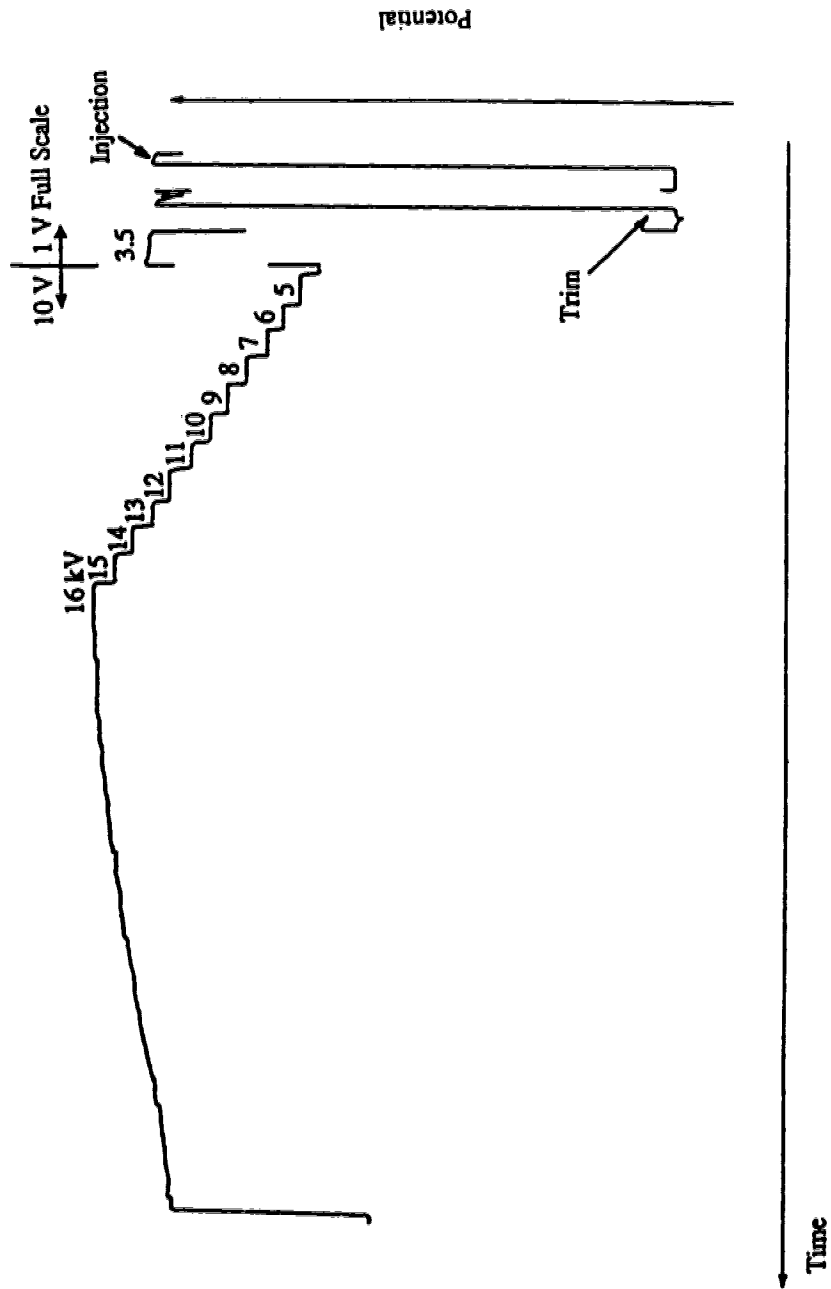


Fig. 2.10. Current profile showing the following stages: 1) sample injection; 2) capillary trimming; 3) stepwise voltage increase and 4) a slow decrease in current during electrophoresis. Each step was 1 minute (except 2 minutes at 3.5 kV) and the whole process took about 38 minutes. There was a scale change from 1 to 10 V full scale.

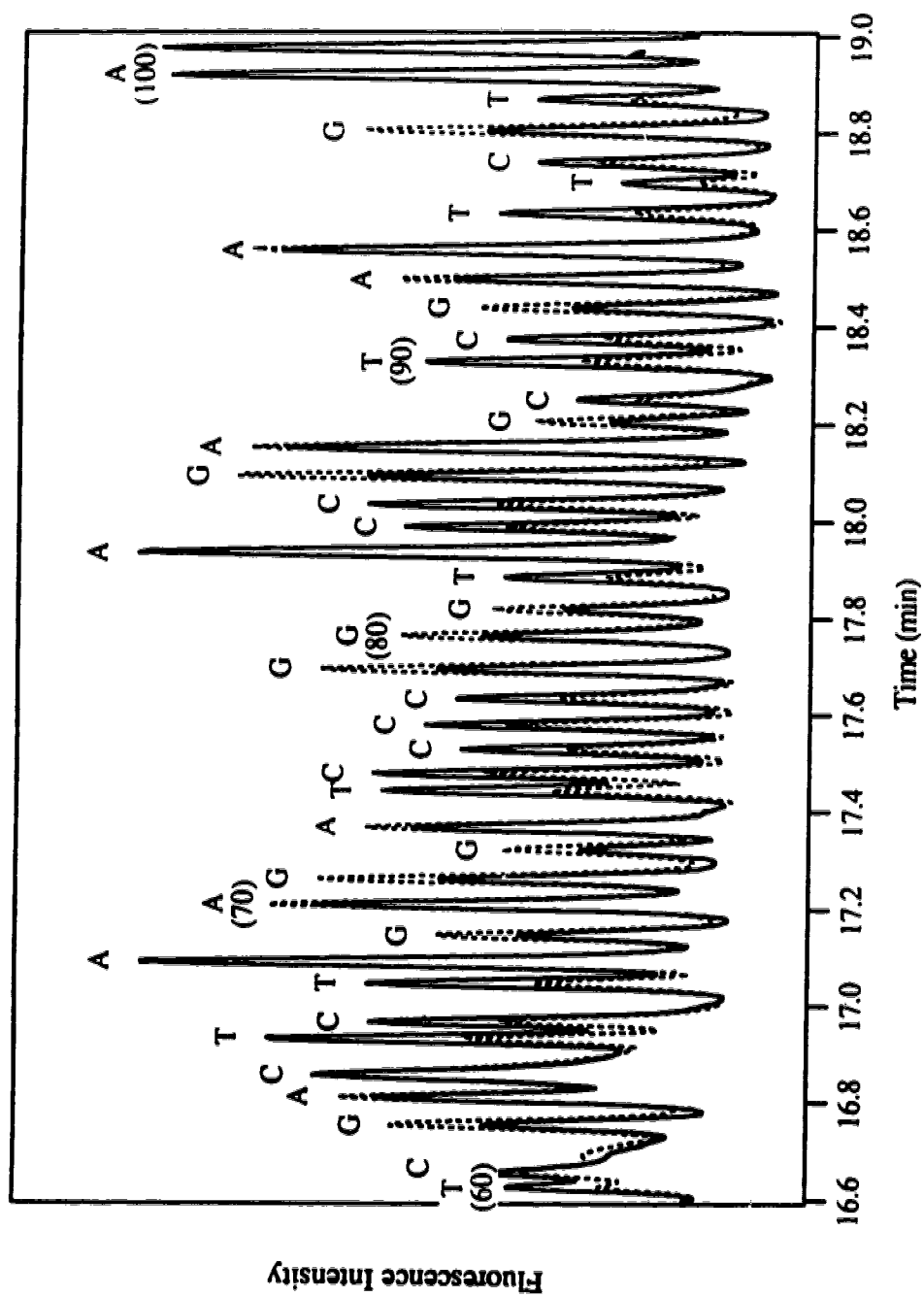


Fig. 2.11. Expanded region of Fig. 2.8a, corresponding to nucleotides 60 - 100 with the known sequence printed above.

fluorescent signal detected in the blue spectral channel centered at 510 nm. In this system, T generates peaks that are much stronger in the red spectral channel compared with the blue channel, C generates peaks for which the red channel amplitude is slightly larger than the blue channel, A generates peaks with roughly equal amplitude in the two channels, and G generates peaks for which the blue channel amplitude is larger than the red channel.

Careful inspection of the data reveals that the accuracy of the sequence data is not high; several errors can be noted in the data, particularly bases 64 and 83, which produce similar peak ratios, but are associated with C and A, respectively. The DuPont sequencing protocol suffers from limited accuracy because a relatively small difference in the ratio of peak amplitude is used to identify each peak. Noise corrupts sequence identification.

Prober *et al.* have reported that at high angles of incidence, the passband of the filter is shifted toward the excitation wavelength, increasing background levels and associated noise [20]. They have used a fiber-optic face plate with an extramural absorber to reduce this effect. However, fluorescence polarization may present another problem that has been previously overlooked.

When fluorescent molecules are excited with polarized light, those molecules with their absorption dipole aligned with the electric vector of the polarized light have the highest probability of excitation and the resulting fluorescence is also polarized. But during the fluorescence lifetime, these molecules are free to rotate, resulting in depolarization. The degree of fluorescence polarization, P , depends on the size of the molecule, V , on viscosity of the solvent, η , and on temperature, T [26]:

$$\frac{1}{P} = a \left(1 + \frac{\tau \cdot kT}{\eta \cdot V} \right) \quad (2-3)$$

where a is a constant. Generally speaking, larger molecules and more viscous

environments will generate more polarized fluorescence.

It is interesting to note that in DNA sequencing, we have a wide range of sizes of DNA fragments and we can imagine that fluorescence becomes more and more polarized during the process of electrophoretic separation when the size of the DNA fragments becomes larger and larger, particularly for on-column detection where the gel provides a more viscous environment than buffer solution. The fact is that the cutoff wavelength of the dichroic beamsplitter is red-shifted by about 5 nm when the source is totally polarized (Fig. 2.12). Because DuPont's dyes have closely spaced emission spectra, any slight shift in wavelength may lead to an adverse effect on the ratio of fluorescence intensity, and so, to the sequence determination, although this effect may be overwhelmed by the system noise.

An alternative dye-terminator system developed by ABI uses four fluorescein labels [27] (see Fig. 1.4 in Chapter 1) with relatively wider absorbance and emission spectra and is promising for DNA sequencing, although two laser lines and four spectral channels are required in its application. The result with this dye system will be presented in Chapter 4.

It should be noted that the two-spectral channel instrument presented in this chapter has proven to be very useful in DNA sequence determination by using the two-color, peak-height approach (see section 1.3.6 in Chapter 1), where the two spectral channels are spaced by 20 to 30 nm. DNA sequence determination accuracy up to 99.5 % has been achieved for the first 400 bases by Chen *et al.* in this group [28].

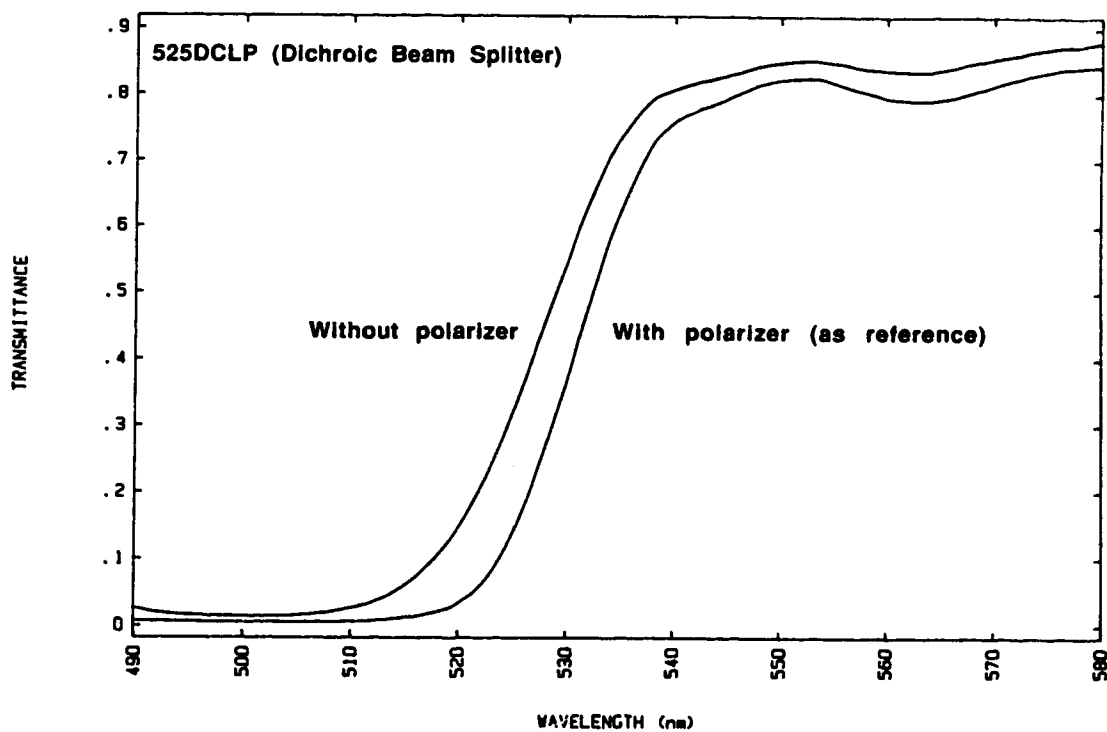


Fig. 2.12. The effect of polarized light on the spectrum of the dichroic beam splitter, 525DCLP. The polarized light was achieved by putting a dichroic sheet polarizer, 03 FPG001 (Melles Griot), in front of the light source of a Diode Array Spectrophotometer (Model 8451A, Hewlett Parckard).

REFERENCES

1. Edström, J.E., *Nature*, 1953. **172** : p. 809.
2. Edström, J.E., *Biochim. Biophys. Acta* , 1956. **22** : p. 378-88.
3. Hjertén, S., *J. Chromatography*, 1983. **270** : p. 1-6.
4. Cohen, A.S. and Karger, B.L., *J. Chromatography*, 1987. **397** : p. 409-417.
5. Cohen, A.S., Najarian, D.R., Paulus, A., Guttman, A., Smith, J.A., and Karger, B.L., *Proc. Natl. Acad. Sci. USA* , 1988. **85** : p. 9660-9663.
6. Guttman, A., Cohen, A.S., Heiger, D.N., and Karger, B.L., *Anal. Chem.*, 1990. **62** : p. 137.
7. Paulus, A., Gassmann, E., and Field, M.J., *Electrophoresis*, 1990. **11** : p. 702-708.
8. Yin, H.-F., Lux, J.A., and Schomburg, G., *J. High Resolution Chromatography*, 1990. **13** : p. 624-627.
9. Bente, P.F. and Myerson, *U. S. Pat.*, 1989. **4,810,456** .
10. Karger, B.L. and Cohen, A.S., *U.S. Pat.*, 1989. **4,865,706** .
11. Swerdlow, H. and Gesteland, R., *Nucleic Acids Research*, 1990. **18** : p. 1415-1419.
12. Drossman, H., Luckey, J.A., Kostichka, A.J., D'Cunha, J., and Smith, L.M., *Anal. Chem.*, 1990. **62** : p. 900-903.
13. Luckey, J.A., Drossman, H., Kostichka, A.J., Mead, D.A., D'Cunha, J., Norris, T.B., and Smith, L.M., *Nucleic Acids REsearch* , 1990. **18** (15): p. 4417-4421.
14. Cohen, A.S., Najarian, D.R., and Karger, B.L., *J. Chromatography*, 1990. **516** : p. 49-60.
15. Karger, A.E., Harris, J.M., and Gesteland, R.F., *Nucleic Acids Research*, 1991. **19** (18): p. 4955-4962.
16. Swerdlow, H., Wu, S., Harke, H., and Dovichi, N.J., *J. Chromatography*, 1990. **56** : p. 61-67.

17. Swerdlow, H., Zhang, J.Z., Chen, D.Y., Harke, h.R., Grey, R., Wu, S., Dovichi, N.J., and Fuller, C., *Anal. Chem.*, 1991. **63** : p. 2835-2841.
18. Chen, D.Y., Swerdlow, H.P., Harke, H.R., Zhang, J.Z., and Dovichi, N.J., *J. Chromatography*, 1991. **559**: p. 237-246.
19. Chen, D. and Dovichi, N.J., *Unpublished data*.
20. Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A., and Baumeister, K., *Science*, 1987. **238** : p. 336-341.
21. Knoll, J.E., *J. Chromatography*, 1985. **23** : p. 422-425.
22. Huang, X., Gordon, M.J., and Zare, R.N., *Anal. Chem.*, 1988. **60** (4): p. 375.
23. Wu, S. and Dovichi, N.J., *Talanta*, 1992. **39** (2): p. 173-178.
24. Nishkawa, T. and Kambara, H., *Electrophoresis*, 1991. **12** : p. 623-631.
25. Swerdlow, H., Dew-Jager, K.E., Brady, K., Grey, R., Dovichi, N.J., and Gesteland, R., *Electrophoresis*, 1992. **13** : p. 475-483.
26. Parker, C.A., *Photoluminescence of solutions*. 1968, New York: Elsevier. 51-63.
27. Lee, L.G., Connell, C.R., Woo, S.L., D., C.R., McArdle, B.F., Fuller, C.W., Halloran, N.D., and Wilson, R.K., *Nucleic Acids Research*, 1992. **20** (10): p. 2471-2483.
28. Chen, D., Harke, H.R., and Dovichi, N.J., *Nucleic Acids Research* , 1992. **20** (18): p. 4873-4880.

Chapter 3

Multiple Capillary System (1): A preliminary Investigation

3.1. INTRODUCTION

Large-scale DNA sequencing [1] requires instruments that generate high throughput and high accuracy at low cost. It is necessary that either novel approaches be fully developed (including time-of flight mass spectrometry to separate Sanger sequencing fragments [2], hybridization to a plate containing a complete set of hexamers [3], scanning tunneling microscopy of an individual strand of DNA [4], and single molecule detection of nucleotides cleaved from a single strand of DNA [5]), or that the conventional technologies be pushed to their limits. The success of capillary gel electrophoresis leads some groups to investigate slab gel techniques with ultrathin, 10- to 100- μm -thick, gels on optically polished glass plates; while other groups have decided to further develop the capillary technique, making use of multiple capillaries that can run many DNA samples simultaneously.

Kostichka *et al.* reported the use of thin slab gels for high speed separation of DNA fragments [6]. The thin slab gel has similar heat transport properties to an array of capillaries, which allows the use of high electric fields. In Smith's instrument, 0.6 mm wide lanes with 0.4 mm interlane spacing were employed to separate 18 samples simultaneously in a 75- μm thick slab gel. This system operated at 40 °C and at an electric field strength of roughly 200 V/cm; 450 bases of sequence were determined in one hour of separation time. This system produced detection limits of between 12 and 60 amol of

rhodamine-labeled primer with a CCD as a detector. Stegemann *et al.* reported a similar approach where 100- μm thick slab gels composed of 3%T, 3%C polyacrylamide was used for DNA sequencing under electric fields of 75-84 V/cm, generating a sequencing speed as high as 1200 bases per hour (probably at much higher temperature) [7].

Zagursky and McCormick [8] have reported DNA sequencing separations performed in a capillary-like format. Their instrument was able to accommodate as many as 12 gel-filled tubes with 530- μm ID at one time. The instrument used the four dye-terminator approach of Prober's ([9], described in Chapter 2). Sequence was determined out to 500 bases with 96% accuracy in a separation time of 9.5 hours. This separation was limited by the Joule heating and was performed at low electric field strength.

Other reports on multiple capillaries came from Mathies and Huang in 1992 [10] and Kambara and Takahashi in early 1993 [11]. Mathies used a more sophisticated scanning instrument to image an array of capillaries. Mathies' instrument operated with 100- μm ID capillaries and produced sequencing information up to 320 bases in length. The Hitachi instrument was based on a sheath flow cuvette fluorescence detector. The instrument has only separated the products of single base terminated reaction. Also this year, Yeung reported a fiber-optic array to excite fluorescence from a set of capillaries [12]. That system has not yet been used for DNA sequencing.

This group has many years of experience with capillary electrophoresis, with the laser-induced fluorescence detection and with the use of the sheath flow cuvette as a fluorescence detector [13-16]. The instrument described here is based on a modified sheath flow cuvette as a fluorescence detector for multiple capillaries. Instrument performance was tested in the form of multiple open tube capillaries and multiple gel-filled capillaries. The results shown in this chapter, although rather preliminary, demonstrate the feasibility of the multiple capillary system as a separation medium based on the sheath flow cuvette as a fluorescence detector.

3.2. INSTRUMENTAL

3.2.1. DESIGN CONSIDERATION

General

To achieve higher sequencing rate, a multiple capillary instrument will simultaneously separate many samples on multiple gel-filled capillaries. These capillaries will terminate in a modified rectangular sheath flow cuvette. In this cuvette, capillaries are aligned like the teeth of a comb. The sheath flow draws the samples from each capillary in a separate stream to form parallel streams in the cuvette. Fluorescence from all sample streams is excited simultaneously with a single (or two) laser line(s) along the long axis of the cuvette (Fig. 3.1), collected by a high numerical aperture microscope objective, filtered by a (or a set of) carefully chosen interference filter(s), and detected by a set of photodetectors. A single high voltage power supply will drive the electrophoresis and a single computer will collect, analyze, and store data.

Capillary quality assurance

The quality of the gel-filled capillaries is another important issue for the multiple capillary instrument. A very long capillary can be poured and then cut to size; or many short capillary gels can be poured separately. In any case, the capillary gels should be bubble-free.

The presence of bubbles in the gel will lead to degradation of separation efficiency of the capillary. In particular, small bubbles may act as sites where catastrophic breakdown of the gel occurs. Bubbles can be detected optically. A laser-light scatter apparatus can be constructed to automatically inspect arbitrary lengths of gel-filled capillary much like the device developed for particle size analysis [17]. However, for this work, a microscope was employed to inspect visually the length of the capillary.

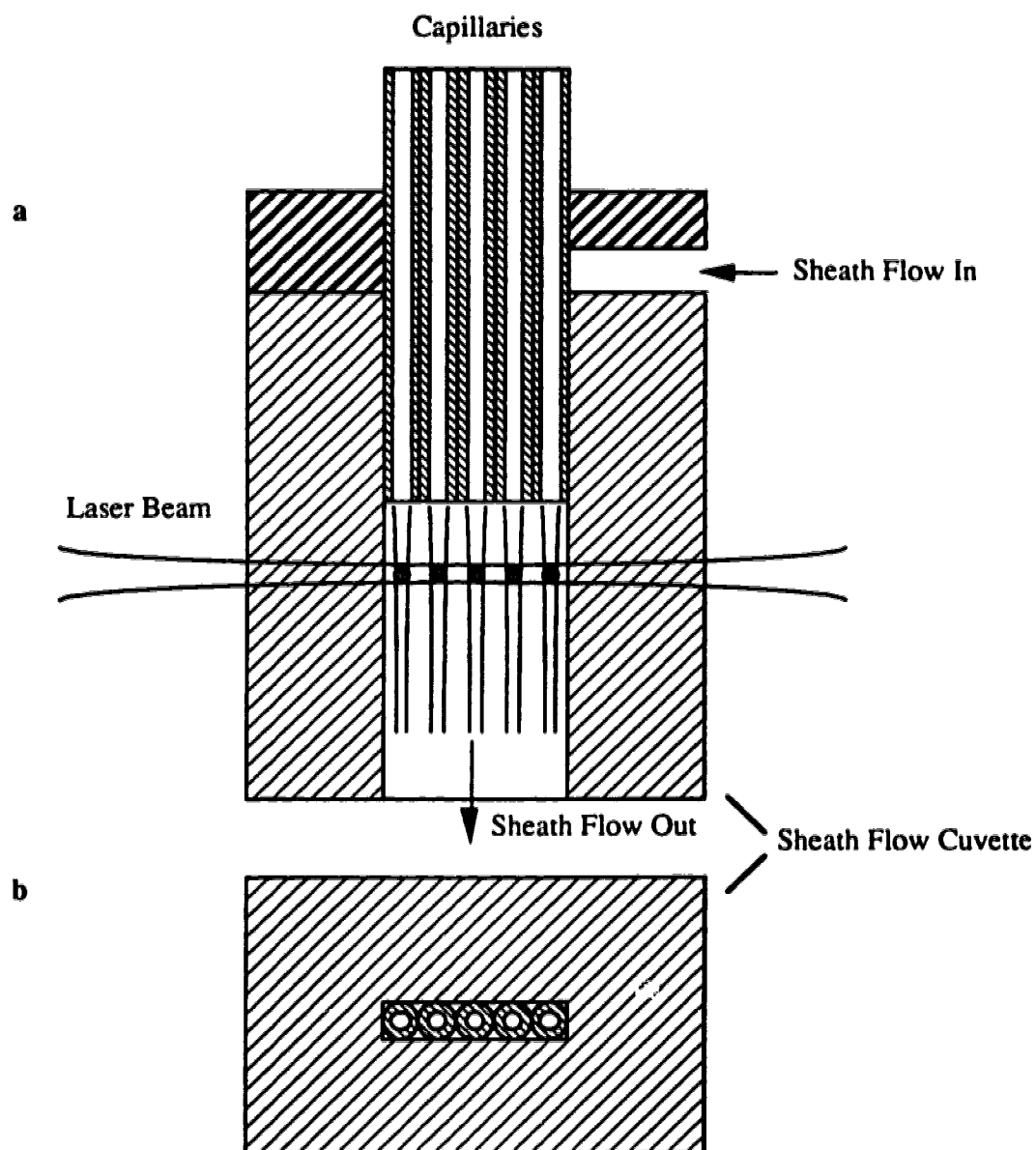


Fig. 3.1. A schematic of sheath flow cuvette based fluorescence detection chamber for an array of five capillaries. The sample streams are isolated by the sheath flow. The laser beam induces fluorescence from these sample streams to form discrete fluorescent spots. Part a is the front view and part b is the top view.

Power supply

There is no severe requirement on the power supply for the 5-capillary instrument. Each capillary has a platinum electrode and all electrodes are connected to a single 30 kV power supply. The sheath flow cuvette is grounded through a current-sampling resistor to complete the electric circuit. Each capillary, 30 to 40 cm long, draws less than 10 μ A current when 1 \times TBE is used as the separation buffer at 15 kV. A 30 W power supply can drive electrophoresis in 200 capillaries simultaneously!

Sample loading

DNA sequencing samples are usually prepared in a volume of several microliters. To ensure proper sample injection, capillaries and electrodes are necessarily trimmed to the same height, or each sample vial can be adjusted to the individual capillary and electrode because there is no constraint on the height of the injection end in gel-filled capillary electrophoresis. The individual height adjustment can be achieved by placing a small piece of spring or sponge under the sample vials.

Sample

The multiple capillary instrument should be designed for a wide variety of DNA sequencing samples, such as one-label, peak-height encoded sample (see Section 1.3.5. one-label peak-height), two-label, peak-height (see Section 1.3.6. two-label peak-height), four-dye, four-spectral-channel sample (see Section 1.3.2. four-dye primers and Section 1.3.3. four dye-terminators). Of course, one or two laser line(s) and a minimum number of optical elements should be used for the instrument.

Fluorescence detection

In fluorescence detection, it is necessary simultaneously to maximize the fluorescence signal and to minimize the noise in the background signal. An on-column

detector suffers from high background signal generated from the light scatter at the capillary-air and capillary-solution interfaces and from the gel itself. Furthermore, introduction of laser light to a large number of capillaries is a difficult task, requiring individual illumination beam for each capillary.

An interesting design choice exists for the collection optic and photodetector. To avoid the use and alignment of pinholes, the magnification provided by the collection optic must provide an image that is matched in size to the photodetector. The excellent image produced by a highly corrected, flat-field, high numerical aperture microscope objective is ideally suited to application with the multiple capillary system. Alternative collection optics, such as a simple lens or elliptical mirror, either have relatively poor collection efficiency or are incapable of producing an image. Assuming a sample stream diameter of 50 μm and a photodetector diameter of 1 mm, the collection optic needs to produce a 20 \times image of the illuminated sample stream. It should be noted that a large array of capillaries would expand the field of view, and a wider field of view would require use of lower power objectives.

Unfortunately, the use of a microscope objective for the collection optic prohibits the use of some photodetectors, such as image-intensified diode-arrays, or charge coupled devices (CCDs), simply because the detection element is much smaller than that produced by well corrected, high numerical aperture microscope objective. Other detectors, such as PMT, produce excellent amplification and reasonable quantum yield but require the use of a pinhole and tend to be bulky and expensive. The PMT would not be a practical photodetector in a multiplexed instrument.

Instead, the RCA C30902S silicon avalanche photodiode (SiAPD) is an interesting choice as the photodetector. This solid state device has high quantum efficiency—50% at 550 nm. It can be operated in the Geiger mode, which provides photon counting, or in the analog mode, which provides sub-nanosecond response. The former mode will be useful if the background signal is low whereas the latter mode may

be used when background count is high. The only thing that restricts its use, however, is its relative small sensitive area.

3.2.2. SILICON AVALANCHE PHOTODIODE AND GRADIENT INDEX LENS

SiAPD

For ultra-low-light detection in the visible range, the designer can choose either a PMT or SiAPD as the single-element detector, because both devices can provide internal gain [18]. Table 3.1 compares some important parameters of two high-sensitivity photodetectors, Hamamatsu R1477 PMT and RCA C30902S SiAPD.

Like a PMT, the SiAPD also works under high biased voltage; the high reverse bias voltage leads to a high field in the p-n junction region. Any carrier, photo-generated or thermal-generated, can be accelerated in this region to an energy level at which collisional ionization occurs. This process leads to carrier amplification or avalanching,

Table 3.1 Parameters relevant to PMT and SiAPD

Parameters	PMT	SiAPD
Active diameter (mm)	~ 10	~ 0.5
Bias requirements	600 to 1200 V	180 to 250 V
Spectral range (nm)	185 to 900	350 to 1150
Quantum efficiency (%)	~18 @ 500 nm	~ 50 @ 550 nm
	~13 @ 633 nm	~ 60 @ 633 nm
Internal Gain	10^6	250 in analog mode 10^8 in Geiger mode
Lifetime	use dependent	> 10 years

and hence the device provides internal gain. SiAPD can be operated in two modes: analog mode, where the reverse bias is set below the breakdown voltage, and Geiger mode, where the reverse bias is set above the breakdown voltage. EG&G produced a photon counting module (SPCM-100-PQ, EG&G Canada Ltd., Québec, Canada), in which the SiAPD is operated in Geiger mode. This photon counting module needs a quenching circuit for the SiAPD (Fig. 3. 2). The principle is that the SiAPD must maintain a current higher than a particular level $\sim 50 \mu\text{A}$ for C30902S) in order to be conducting or it can switch off by itself. The circuit is called passive-quenching. In this passive-quenching mode, the photodetector has a dead time of 200 ns between pulses. During the dead time period, the SiAPD can not detect the photons that reach it. As a result, the true count is slightly greater than the observed count. This passive-quenching circuit could let the counting rate be as high as 1 to 2 MHz. Above this rate, the SiAPD is saturated.

The SiAPD's operation is very temperature-sensitive; the breakdown voltage, V_R , has a typical temperature coefficient of $0.7 \text{ V}/^\circ\text{C}$. Either a feedback circuit or temperature control is necessary for its stable operation. The SPCM-100-PQ is provided with a temperature control unit. Table 3.2 provides some of the electrical characteristics of this device.

It is the virtue of high quantum efficiency plus internal gain that makes it very attractive to the ultra-high sensitivity fluorescence detection. However, its small active diameter limits its use.

In the process of instrument design, we found that an inexpensive optical element, the gradient index (GRIN) lens was very helpful. It is able to focus a collimated beam into a small spot that could match the active diameter of the SiAPD. Also, it eliminates the necessity of pinholes, as its size acts as a pinhole by itself.

GRIN lens

The operational principle of GRIN lens is quite different from that of a

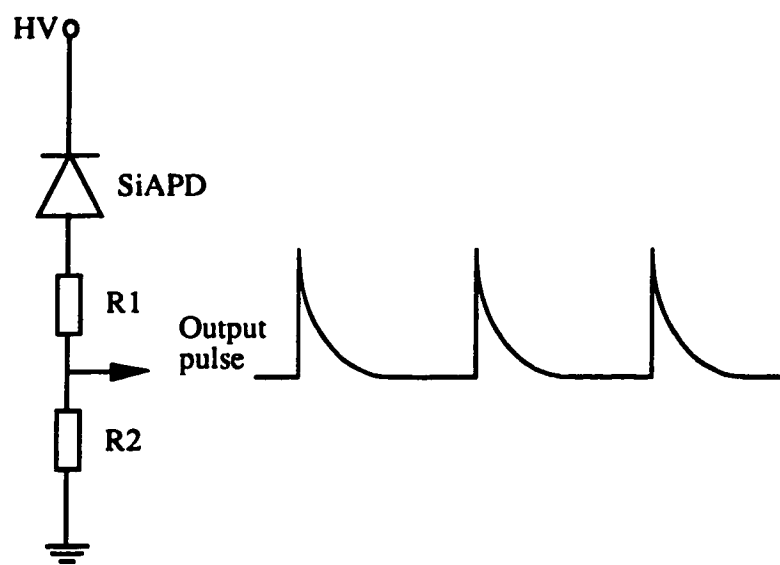


Fig. 3.2. A schematic of passive quenched circuit and the resulting pulses. The load to the SiAPD is the sum of resistors R_1 and R_2 . $R_1 \gg R_2$. R_2 also functions as the sampling resistor where the pulses are generated.

Table 3.2. Data Sheet for SPCM-100-PQ (Selected)

Electrical Characteristics	Minimum	Typical	Maximum	Unit
Photon Detection				
Efficiency Within 50 μm of Center:				
$\lambda = 830 \text{ nm}$	25	32	—	%
$\lambda = 633 \text{ nm}$	40	46	—	%
$\lambda = 400 \text{ nm}$	—	15	—	%
Operating Temperature	—	0	—	$^{\circ}\text{C}$
Dark Count	—	250	1000	counts/s
Dead Time	—	200	—	ns
Output Count Rate Before Saturation	1.3×10^6	1.8×10^6	—	counts/s
Single Photon Timing				
Resolution (rms)	—	3	—	ns
Settling Time Following Power Up	—	20	100	s
Threshold Setting Required for Digital Output Pulse (Terminate in 50 ohms)	0.5	1.0	1.5	V

conventional lens. A conventional lens consists of a homogeneous material with a single refractive index. An image is produced by means of a series of discrete refraction at the lens surface-air (or any neighboring material) interfaces based on Snell's law. Image formation can be shown in Fig. 3.3a. A GRIN lens is made of gradient index material that refracts the light continuously, as well as discretely at the surface. A GRIN lens can be made into a rod shape where the refractive index, N , is given by

$$N(r) = N_0 \left(1 - \frac{A}{2} r^2 \right) \quad (3-1)$$

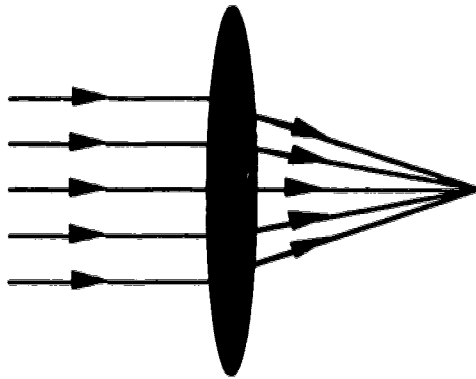
where r is the radius of the GRIN rod lens, N_0 is the refractive index at $r = 0$, and A is the gradient constant. A schematic of a GRIN rod lens is shown in Fig. 3.3b along with the refractive index profile.

A GRIN lens is typically one to several mm in diameter. Light travels through the lens in a sinusoidal fashion. One pitch (P) is defined as the length of the rod required for light of a certain wavelength to execute one cycle, or sine wave. A GRIN rod with $0.25 P$ is able to transform a collimated beam into a point spot (or visa versa) as shown in Fig. 3.3b. This compact optical element offers design flexibility in our fluorescence detection using SiAPD as the photodetector.

3.2.3. INSTRUMENT

The key part of the instrument was a modified sheath flow cuvette. It was used as the detection chamber for the multiple capillary system. Two sheath flow cuvettes (Precision Cells, NY, USA), which were made of high quality quartz, were constructed for this experiment. One had a chamber that was 200- μm deep and 1000- μm wide in cross section, 2-cm long, and had 2-mm thick windows. This cuvette was designed to house five capillaries, 76- μm ID and 200- μm OD. The other cuvette had a chamber that

a. Conventional Lens



b. GRIN lens

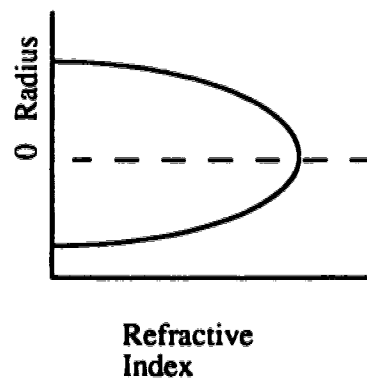
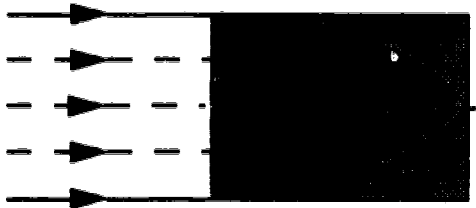


Fig. 3.3. Comparison of a GRIN rod lens with a conventional lens. For a conventional lens (a), light refraction takes place discretely at the air-lens interfaces whereas for a GRIN lens (b), light refracts continuously in the GRIN lens. The refractive index of a GRIN lens is a function of the rod radius.

was 150- μm deep and 750- μm wide in cross section, 2-cm long, and had 1-mm thick windows. This cuvette was also designed to house five capillaries, 50- μm ID and 150- μm OD. Sheath flow buffer, which was identical to the separation buffer, was supplied by a high precision syringe pump, model 906 (Harvard Apparatus, MA, USA) with a flow rate of 1 ml/hour or lower. Sample streams from all capillaries were swept by the sheath flow buffer into the detection chamber like the teeth of a comb. The beam of an argon ion laser (model 2011-20SL, Uniphase, CA, USA), operated at 488 nm in light-regulated mode, was focused with a 2.5 \times microscope objective (Melles Griot, CA, USA) into the sheath flow cuvette, along the cuvette's long axis about 200 μm below the exit end of the capillaries. Fluorescence of all sample streams was excited simultaneously by the focused laser beam, collected at a right angle with a microscope objective, 18 \times , 0.45 numerical aperture (Melles Griot, CA, USA), filtered with a bandpass filter (Omega, VT, USA), imaged onto an array of GRIN lenses (NSG, Tokyo, Japan), transported through array of fiber optics and detected by an array of SiAPDs, or photon counting modules (EG&G, Montreal, Canada) as shown in Fig. 3.4.

Data acquisition was performed by using an NB-MIO-16X board from National Instruments (TX, USA). This multi-purpose board provides three 16-bit counters for counting operation. Because of this limit, in this five capillary instrument, only three separation channels were detected. In the following chapter, data from all five capillaries will be presented by using another data acquisition board, NB-TIO-10, which provides ten 16-bit counters.

The 16-bit counters can be diagrammed as shown in Fig. 3.5. A counter usually has three terminals: source, gate and out. In this experiment, the source signal is the pulses generated from the photon counting module and the gate signal can be either provided by software or by a pulse with certain duration from an external circuit. In this experiment, the gate signal was provided by software. The output is the count value up to 65536 (16 bits). The data acquisition and storage were controlled by a Macintosh IIsi

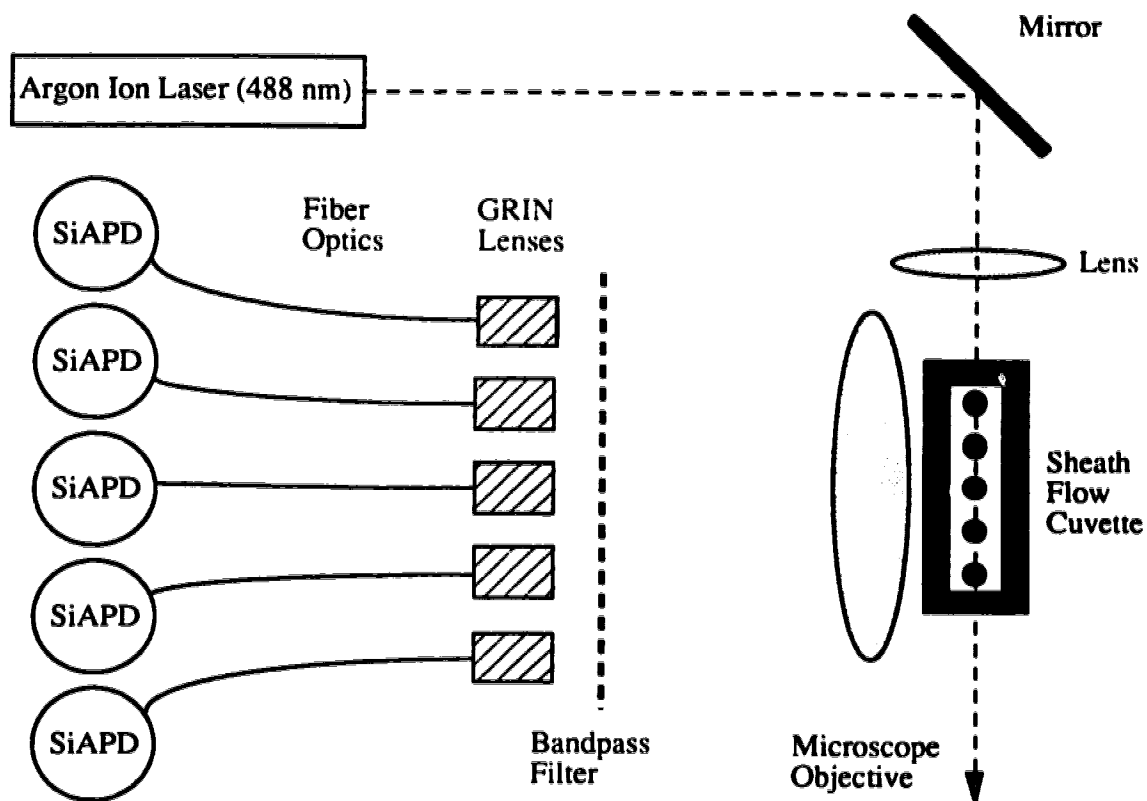


Fig. 3.4. The fluorescence detection system for the multiple capillaries. The discrete sample fluorescence spots are imaged by a single microscope objective onto an array of GRIN lenses. These spectrally filtered fluorescence signals are transported through the fiber optics and detected by an array of discrete SiAPDs.

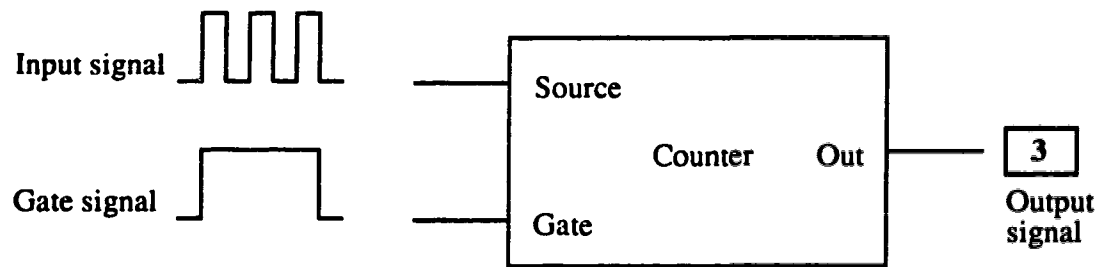


Fig. 3.5. A counter block diagram and the signal forms for source, gate and out terminals.

computer.

3.3. EXPERIMENTAL

3.3.1. REAGENTS

Fluorescein isothiocyanate (FITC), isomer I and amino acids were obtained from Sigma (St. Louis, MO, USA). FITC was dissolved in acetone to a concentration of about 10^{-3} M in a 1.5 ml centrifuge tube and stored in 4 °C. The separation buffer was 10 mM borate buffer, pH 9.2, for amino acids and 1 × TBE, pH 8.3, for DNA fragments.

3.3.2. METHODS

Fifteen FITC-labeled amino acids were prepared similar to the method described in [13]. Briefly, each stock solution of amino acid, 5.0×10^{-3} M in concentration, reacted with FITC, 1.2×10^{-4} M in concentration, for 4 hours in the dark in a 0.2 M borate buffer, pH 9.2. These FITC-labeled amino acids were diluted to a final concentration (FITC) of 1.2×10^{-10} M before separation by free zone electrophoresis.

Gel-filled capillaries were prepared the same way as described in Section 2.3.2, CAPILLARY GEL PREPARATION.

The DNA sequencing sample was prepared according to the procedure described by Richardson and Tabor [19] (see section 1.3.4 in Chapter 1). Briefly, FAM (a nickname of fluorescein) -labeled primer, 8 pmol, was annealed to 10 µg of M13mp18 single-stranded DNA template at 65 °C for 2 minutes followed by slow cooling to ≤ 30 °C. A mixture of deoxynucleotide triphosphates and dideoxynucleotide triphosphates (concentration ratio of 300:1) was added with the ratio of nucleotides adjusted to yield a nominal peak height ratio of 8:4:2:1 for G:T:A:C. The mixture was warmed at 37 °C and

6 units of Sequenase Version 2.0 and 0.006 units of pyrophosphatase were added. Incubation continued at 37 °C for 10 minutes, after which the reaction was stopped and DNA fragments were ethanol precipitated and washed. Finally the DNA sample was resuspended in 4 µl of a 49:1 mixture formamide/0.5 M EDTA solution at pH 8.0. The sample was heated at 95 °C for two minutes before loading into the capillaries.

3.4. RESULTS AND DISCUSSION

3.4.1. DETECTION LIMITS

Linear calibration curves were obtained for FITC for concentrations from 1.0×10^{-11} to 5.0×10^{-10} M in 10 mM borate buffer, pH 9.2 (Fig. 3. 6). This buffer was also used as running buffer and sheath flow fluid. Five capillaries, 76 µm ID, 194 µm OD and 46.7 cm long, were inserted into the sheath flow cuvette chamber, but only three adjacent capillaries were used. FITC was injected hydrodynamically for 10 seconds at a height difference of 20 mm between the water level of the injection end and that of the waste container. Sheath flow was stopped to avoid a back pressure caused by sheath flow [20]. The volume of sample injected, V , is calculated to be 4 nl, according to the following equation:

$$V = \frac{\pi r^4 \rho g t_i}{8 \eta L_c} \Delta h \quad (3-2)$$

where r is the radius of the capillary inner radius, 37 µm, Δh is the height difference, 20 mm, L_c is the length of the capillary, 46.7 cm, g is the acceleration due to gravity, 9.8 m/s², ρ is the density of the sample solution, ~1 g/cm³, and η is the viscosity of the solution, 9548 cP at 22 °C. The slope of the calibration curve was different from capillary to capillary, but the relative differences were no bigger than 5%. The LODs

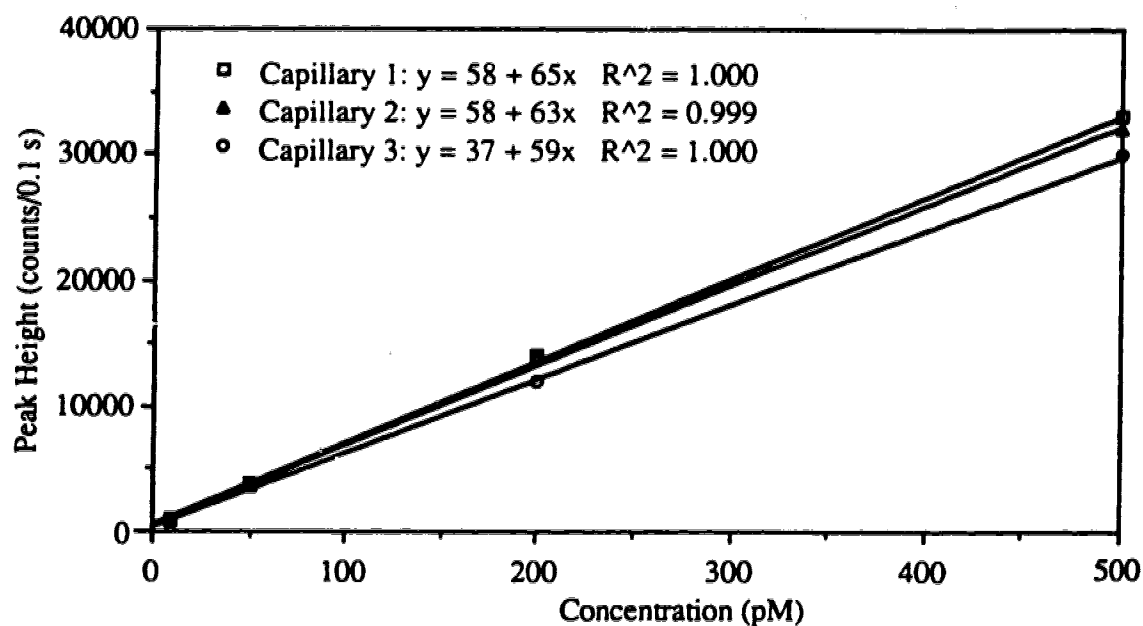


Fig. 3.6. Calibration curves for the three capillary system. Capillaries were 76 μm ID, 194 μm OD and 46.7 cm long. FITC concentrations were 1.0×10^{-11} M, 5.0×10^{-11} M, 1.0×10^{-10} M and 5.0×10^{-10} M. Four nl of FITC solutions were hydrodynamically injected. The sheath flow rate was 1.5 ml/h. The laser power was 7 mW. The peak height was measured as counts per 100 ms.

(3σ) were 20 ± 2 zmol of FITC injected into the capillaries.

Detection limit is an important issue for DNA sequencing by capillary gel electrophoresis. Since the sample loading in capillary gel electrophoresis will be on the order of 1-10 amol of DNA/band [21], detection limits in the zeptomole range are necessary for accurate DNA sequence determination by capillary gel electrophoresis. The detection limit of 20 zmol, although quite preliminary, indicates that the system has an adequate sensitivity that may be useful for DNA sequencing by capillary gel electrophoresis.

One limitation with the photodetector operated in Geiger mode is the limited linear range, about 2 orders of magnitude in FITC concentration. This limited range indicates that for amino acid and DNA fragment separations, suitable sample concentration or injection time must be chosen so that the signals are within the limited linear range of the detector.

3.4.2. FLUORESCCEIN LABELED AMINO ACID SEPARATION BY CAPILLARY ZONE ELECTROPHORESIS

Fig. 3. 7 illustrates the separation and detection of three groups of FITC-labeled amino acids. In each group, there were five FITC-labeled amino acids. The capillaries were the same as used for the detection limit measurements. The injection ends of the three adjacent capillaries were separately inserted into three different buffer or sample vials. The separation buffer and sheath flow fluid were 10 mM borate buffer, pH 9.2. The sheath flow rate was 1.5 ml/h. The samples were hydrodynamically injected for 10 seconds (manually timed) at a height difference of 2 cm, and separation proceeded at 300 V/cm for 7 minutes. The fluorescence signals were sampled at 100 ms time interval by the Macintosh IIsx computer.

Free zone electrophoresis was first applied to this multiple capillary system to

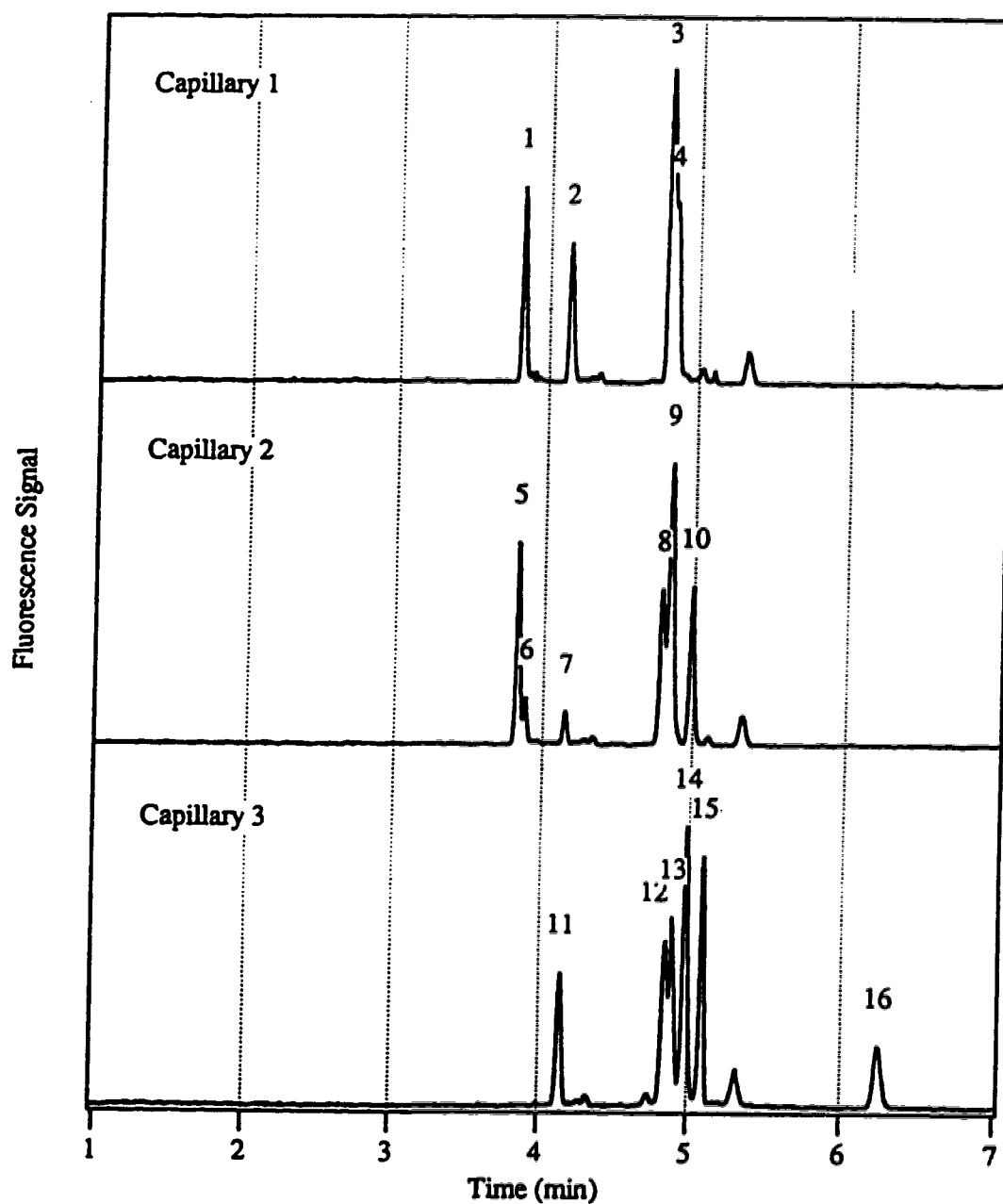


Fig. 3.7. Electropherogram showing simultaneous separation of 3 groups of amino acids. Each group was comprised of 5 amino acids. In capillary 1, 1 is Arg, 2 is Cys and FITC, 3 is Leu and Trp and 4 is Met. In capillary 2, 5 is Lys, 6 is Lys, 7 is FITC, 8 is Iso, 9 is Glu and Pro and 10 is Ala. In capillary 3, 11 is FITC, 12 is Phe, 13 is Thr, 14 is Gly, 15 is Ser and 16 is Asp.

characterize its performance. Crosstalk among the separation channels might be a problem for multiplexed detectors. In the presence of crosstalk, false signals are generated in one channel as the reflection of signals from its adjacent channels. The sheath flow provides a well defined isolation for adjacent sample streams even when capillaries are squeezed side-by-side. Fig. 3. 7 clearly demonstrates that there is no crosstalk between two adjacent capillaries—capillary 3 shows no signal as the reflection of peak 5 in capillary 2 and capillary 2 shows no signal as the reflection of peak 16 in capillary 3.

One electropherogram of FITC is shown in Fig. 3. 8. The system produced similar responses for individual detection channels when the sample, the length of the capillaries and the separation conditions were identical. However, the migration time of FITC was slightly different from capillary to capillary. In this case, for example, the difference is about one peak width. Similar variation in mobility was observed by Yeung without using the sheath flow cuvette; he reported a relative standard deviation for the migration times of fluorescein of 5.5% [12]. This variation in mobility happens because of variation in charge density on the surfaces inside each capillary (and hence the electroosmotic flow rate), variation in temperature, but not variation in capillary ID. In capillary liquid chromatography, the average velocity of analyte in the capillary is proportional to the square of the ID [22] whereas in capillary zone electrophoresis, velocity is ID-independent. Because of the variation in migration times, great care must be taken when using capillary bundles for preparative purpose where only one separation channel was under observation [23]. As will be seen later, the migration-time difference was also found in polyacrylamide gel-filled capillaries, particularly when gels are crosslinked. Because of variation in mobility among capillaries, DNA sequence determination must be conducted in a single capillary instead of four separation channels as in traditional slab gel electrophoresis.

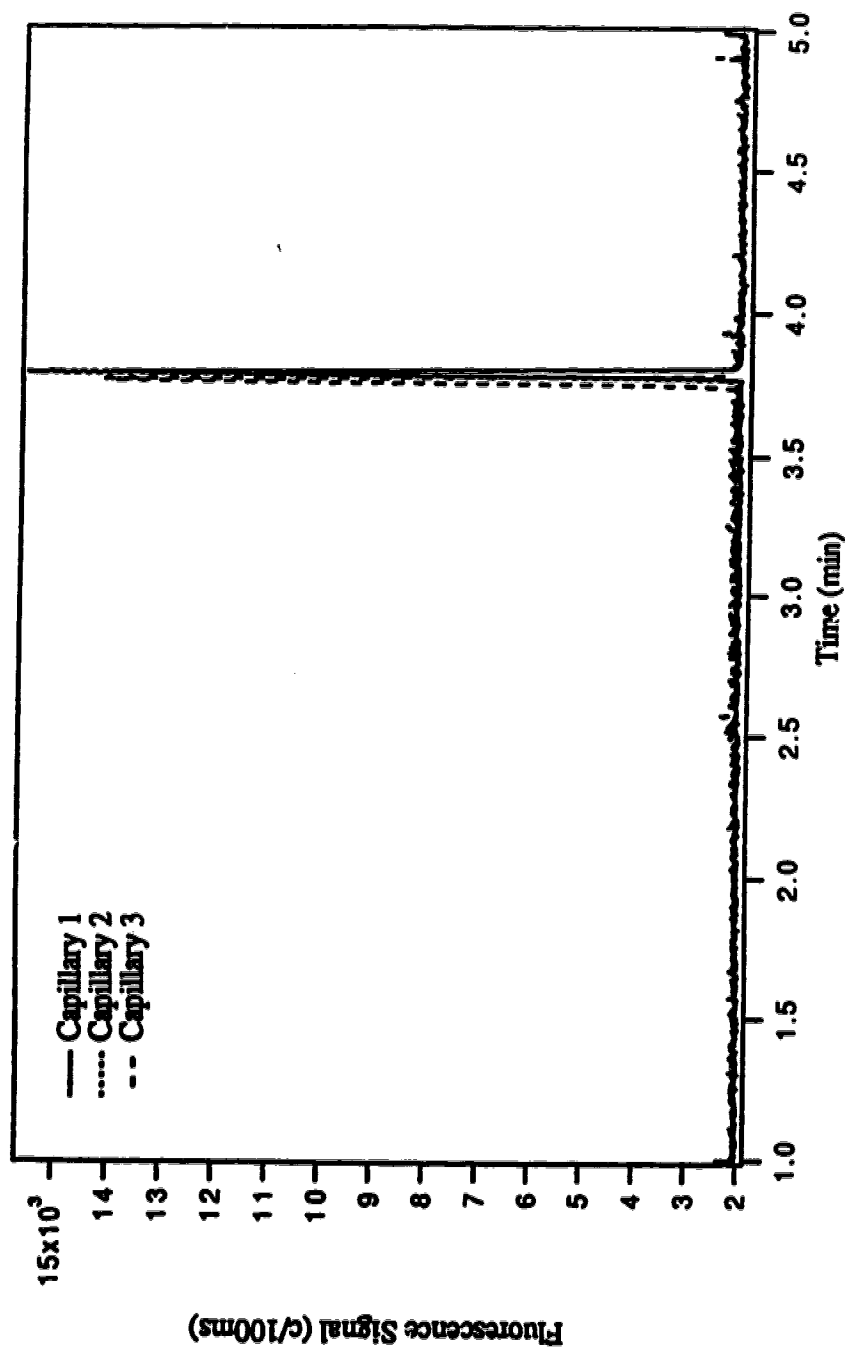


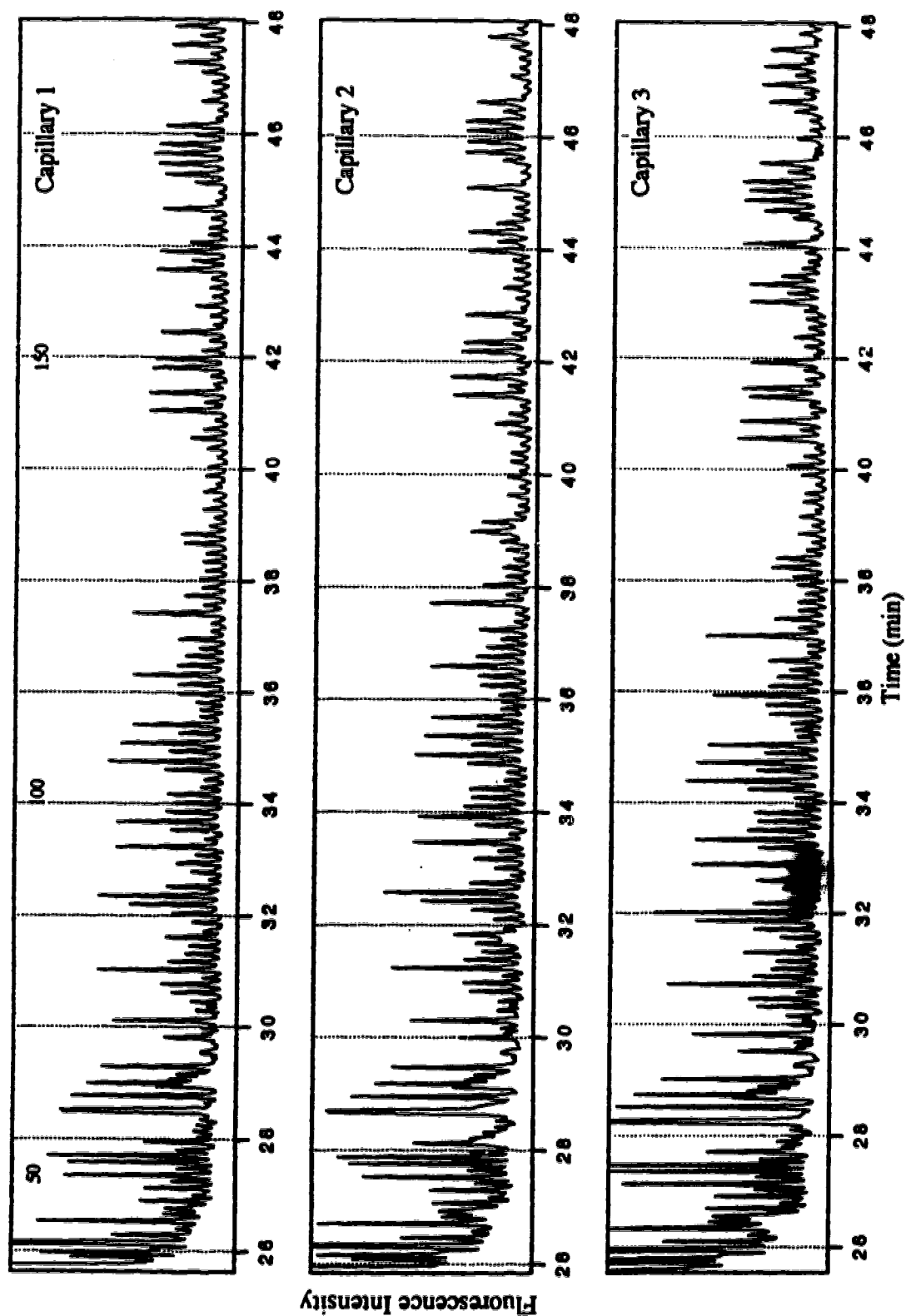
Fig. 3.8. The migration time variation of analyte in three capillaries. The analyte was FITC, 2.0×10^{-10} M in 10 mM borate buffer (pH9.2). The capillaries were 76 μ m ID, 196 μ m OD and 46.7 cm long. Four nl of FITC solutions were hydrodynamically injected. Electrophoresis was conducted at 300 V/cm.

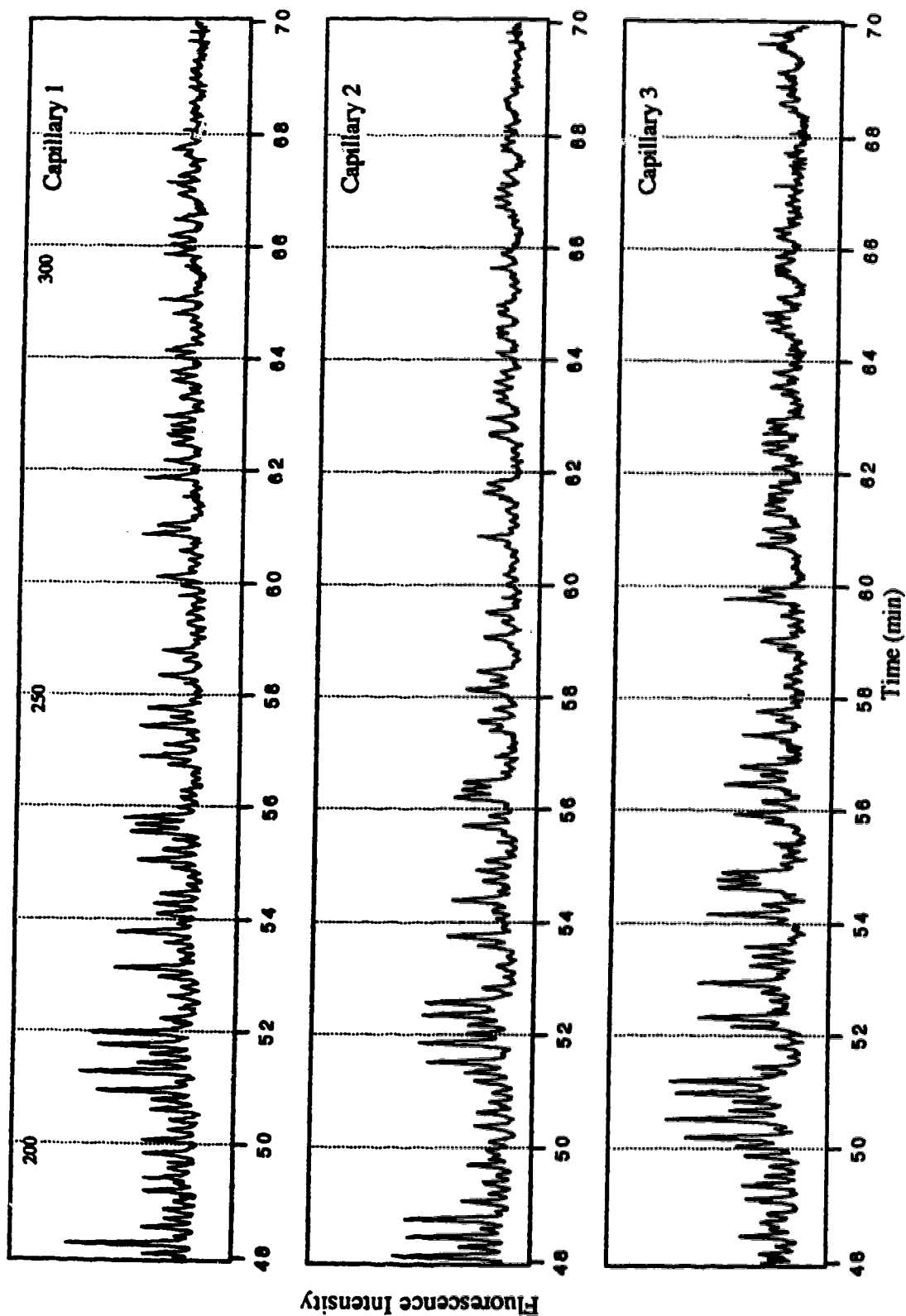
3.4.3. DNA SEQUENCING BY CAPILLARY GEL ELECTROPHORESIS

This multiple capillary system was also tested for DNA sequencing by capillary gel electrophoresis based on a one-spectral-channel technique [16]. The capillaries were filled with 4%T, 5%C polyacrylamide gel with a syringe in a sequential manner. The separation buffer and sheath stream were $1 \times$ TBE. FAM-labeled sequencing sample was prepared according to the method of Tabor and Richardson [19] (see section 1.3.5 in Chapter 1). The sample was injected electrokinetically at 100 V/cm for 60 seconds and electrophoresis was under an electric field of 200 V/cm over 50 μ m ID, 150 μ m OD and 38 cm long capillaries. Fig. 3. 9 shows the result taken with this multiple capillary system. Since the sample was primer-labeled with FAM, the argon ion laser with $\lambda = 488$ nm was also used in this experiment.

The sequencing rate was 400 bases per hour per capillary at a moderate electric field strength of 200 V/cm. Peaks from 50 to 220 bases were nearly baseline resolved except for compression around 60 bases. The accuracy of the sequence determination is low, particularly for small peaks: A and C in many cases are hardly distinguishable. The variation in migration times also occurred in gel filled capillaries as reported by Huang *et al.* [24] and is more pronounced in the report from Kambara and Takahashi [11]. The preparation of many identical gel-filled capillaries for DNA sequencing is very challenging. First, it seems hopeless to fill the gel into a long capillary and then cut it to size. During the process of gel formation, there is a volume change, which will lead to bubble formation along the capillary, particularly when capillaries are treated with silanizing agents [25]. Second, it is difficult to pour many short gels one after another because the gel usually polymerizes within minutes [26-28]. Usually, about five capillary gels can be made in this fashion before the gel is polymerized in the syringe. Gels poured later will have different composition than gels poured first. Third, capillary gels generally are trimmed by a few centimeters at either end, before and/or after sample injection [29].

Fig. 3.9 (pages 89-90). Electropherogram of the three-capillary, one-spectral-channel sequencing of M13mp18. (a) Extended run. The amount of nucleotides were adjusted to yield a nominal peak height ratio of 8:4:2:1 for G:T:A:C. The whole separation process took 70 minutes for fragments up to 330 bases in length. For simplicity, the base number is printed for the first capillary only.





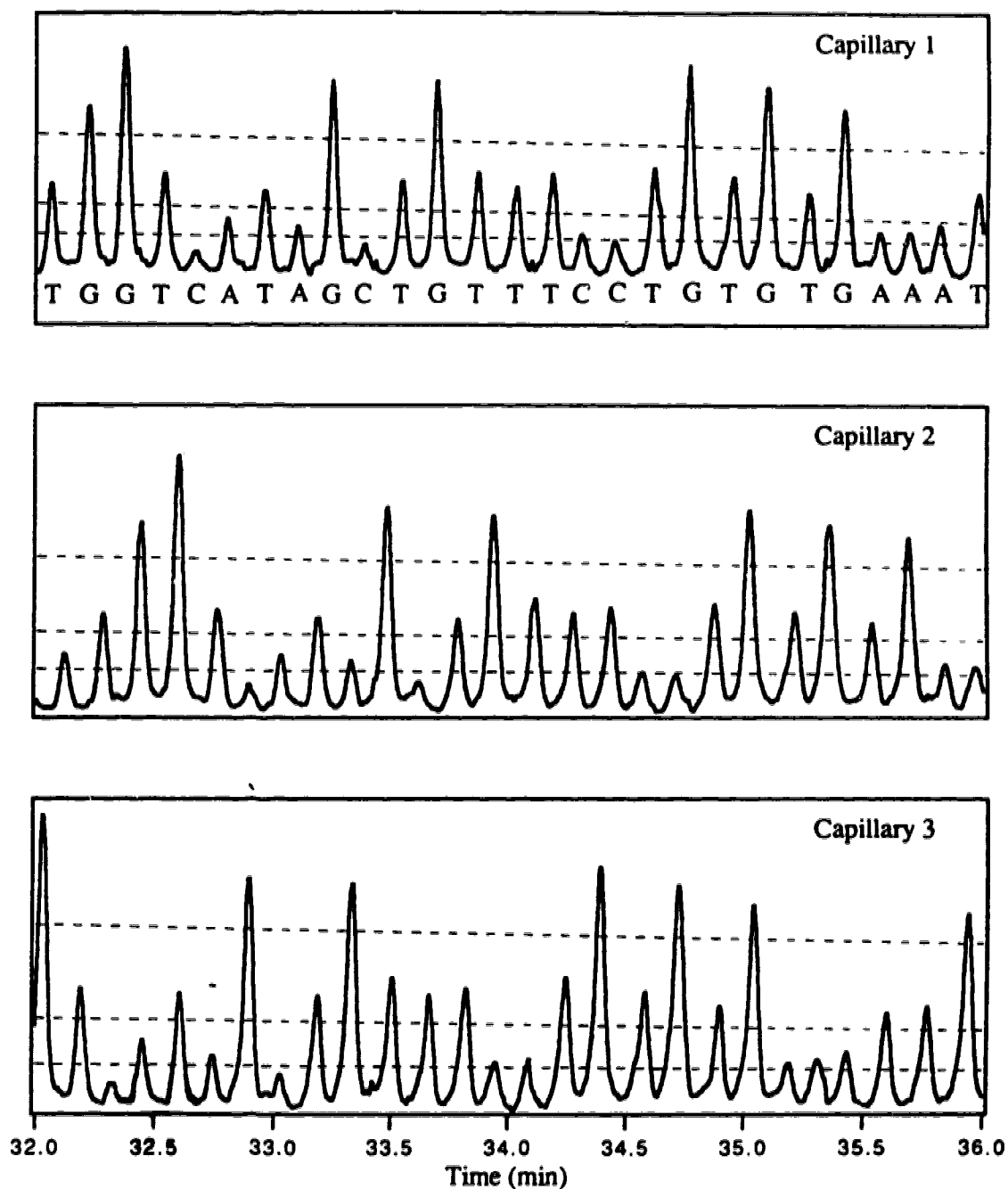


Fig. 3.9. Three-capillary, one-spectral-channel sequencing of M13mp18. (b) expanded region corresponding to 32-36 minutes of the run. Thresholds are set for base identification and the sequence is printed below the peaks for capillary 1.

Trimming can generate bubbles when conventional polyacrylamide gel is been used because of its rigid nature. For these reasons, a new gel matrix must be found and many capillary gels must be poured simultaneously .

3.5. CONCLUSION

The use of multiple capillaries coupled with the sheath flow cuvette fluorescence detector has proven to give a unique combination that provides increased separation speed, high quality isolation of sample streams and adequate detection sensitivity. However, the detection system needs to be improved to gain the sensitivity comparable to that of a single capillary system. Conventional gels need to be replaced in order to make high quality DNA sequencing possible on the multiple capillary system. An improved sequencing technique must be employed to provide high sequencing accuracy.

REFERENCES

1. Hunkapiller, T., Kaiser, R.J., Koop, B.F., and Hood, L., *Science*, 1991. **254** : p. 59-67.
2. Tang, K., *Rapid Comm. mass Spect.*, 1992. **6**: p. 365.
3. Drmanac, R., Labat, I., Brunker, I., and Crkvenjacov, R., *Genomics*, 1989. **4**: 114-128.
4. Allison, D.P., *Proc. Natl. Acad. Sci. (USA)*, 1992. **89** : p. 10129.
5. Davis, L.M., *Genetic Analysis*, 1991. **8** : p. 1.
6. Kostichka, A.J., Marchbanks, M.L., Brumley, R.L.Jr., Drossman, H., and Smith, L.M., *Bio/Technology* , 1992. **10** : p. 78.
7. Stegemann, J., Schwager, C., Erfke, H., Hewitt, N., Voss, H., Zimmermann, J., and Ansorger, W., *Nucleic Acids Research* , 1991. **19** (3): p. 675-676.
8. Zagursky, R.J. and McCormick, R.M., *BioTechniques* , 1990. **9**: p. 74-79.
9. Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A., and Baumeister, K., *Science*, 1987. **238** : p. 336-341.
10. Mathies, R.A. and Huang, X.C., *Nature*, 1992. **359** : p. 167-169.
11. Kambara, H. and Takahashi, S., *Nature*, 1993. **361** : p. 565-566.
12. Taylor, J.A. and Yeung, E.A., *Anal. Chem.*, 1993. **65** : p. 956-960.
13. Cheng, Y.F. and Dovichi, N.J., *Science*, 1988. **242** : p. 562-564.
14. Wu, S. and Dovichi, N.J., *J. Chromatography* 1989. **480** : p. 141-155.
15. Swerdlow, H., Wu, S., Harke, H., and Dovichi, N.J., *J. Chromatography*, 1990. **56** : p. 61-67.
16. Swerdlow, H., Zhang, J.Z., Chen, D.Y., Harke, h.R., Grey, R., Wu, S., Dovichi, N.J., and Fuller, C., *Anal. Chem.*, 1991. **63** : p. 2835-2841.
17. Chen, D., Harke, H.R., and Dovichi, N.J., *Nucleic Acids Research* , 1992. **20** (18):

p. 4873-4880.

18. MacGregor, A., *Photonics Spectra*, 1991. 2: p. 139-146.
19. Tabor, S. and Richardson, C.C., *J. Biol. Chem.*, 1990. 265 : p. 8322-8329.
20. Cheng, Y.F., Wu, S., Chen, D.Y., and Dovichi, N.J., *Anal. Chem.*, 1990. 62 (5): p. 496-503.
21. Drossman, H., Luckey, J.A., Kostichka, A.J., D'Cunha, J., and Smith, L.M., *Anal. Chem.*, 1990. 62 : p. 900-903.
22. Meyer, R.F., Champlin, P.B., and Hartwick, R.A., *J. Chromatogr. Sci.*, 1983. 21 : p. 433-438.
23. Hanai, T., Hatano, H., Nimura, N., and Kinoshita, T., *J. High Resolution Chromatogr.*, 1990. 13 : p. 573-574.
24. Huang, X.C., Quesada, M.A., and Mathies, R.A., *Anal. Chem.*, 1992. 64 : p. 967-972.
25. Yin, H.-F., Lux, J.A., and Schomburg, G., *J. High Resolution Chromatography*, 1990. 13 : p. 624-627.
26. Gelfi, C. and Righetti, G., *Electrophoresis*, 1981. 2: p. 213-219.
27. Gelfi, C. and Righetti, G., *Electrophoresis*, 1981. 2: p. 220-228.
28. Righetti, G., Gelfi, C., and Bosisio, A.B., *Electrophoresis*, 1981. 2: p. 291-295.
29. Swerdlow, H., Dew-Jager, K.E., Brady, K., Grey, R., Dovichi, N.J., and Gesteland, R., *Electrophoresis*, 1992. 13 : p. 475-483.
30. Huang, X.C., Quesada, M.A., and Mathies, R.A., *Anal. Chem.*, 1992. 64 : p. 2149-2154.

Chapter 4

Multiple Capillary System (2):

High performance DNA Sequencer

4.1. INTRODUCTION

For single capillary systems, a simple instrument design is required to maximize the fluorescence collection while minimizing background signal [1]. A similar optical design is required for the multiple capillary system. However, the sensitivity of individual channels of this compact multiple capillary system also relies on the quality of the sheath flow pattern, on the uniformity of excitation of each sample stream, and on the performance of individual optical detection channels. This section will describe some modifications and improvements to the instrument that was described in Chapter 3.

As separation media, gels play an important role in DNA sequencing. Cross-linked gel appears to be suitable only for the use in slab format and in single capillary electrophoresis. It is hard to imagine that cross-linked gels can be successfully applied to multiple capillary systems for the reasons discussed at the end of Chapter 3. Linear acrylamide gel has proven to be very useful for DNA sequencing fragment separation in multiple capillary format [23].

Currently, several fluorescence-labeling strategies have been developed to facilitate automated DNA sequencing (see section 1.3 of Chapter 1). Among those, the dye-terminator method is of particular interest, because only one extension/termination reaction is required to produce DNA sequencing sample and only one lane (or capillary) is required for separation. This method may be used in primer-walking strategies. In this Chapter, the multiple capillary system, with its high sensitivity, is demonstrated in DNA sequence determinations of four-color dye-terminator samples.

4.2. INSTRUMENTAL

4.2.1. SHEATH FLOW PROFILE

As mentioned before, the key to the fluorescence detector is the use of the sheath flow cuvette. The sheath flow draws the sample streams from capillary into the flow chamber. The sample streams should be aligned like the teeth of a comb in the flow chamber. Because the refractive indices of the sheath and sample streams are identical, a single laser beam may be used to illuminate all the sample streams. Because of the minute absorbance of the highly diluted analyte, there is no attenuation of the laser beam in traversing the array of sample streams. Furthermore, the high optical quality of the flow chamber minimizes light scatter, which produces low background signal and excellent detection limits.

However, the sheath flow cuvettes described in the previous chapter were slightly too big to house five capillaries. There was usually a free space at one side of the flow chamber, leading to a non-uniform flow profile across the sheath flow chamber. The non-uniform sheath flow was likely to push the sample streams toward one side of the flow chamber, degrading the flow pattern of the sample streams. The higher the sheath flow rate, the more the degradation. As a result, the sheath flow rate was kept very low, 1 ml/h or lower, to relieve this effect. To eliminate this sheath flow artifact, however, a new sheath flow cuvette was designed.

In the sheath flow cuvette designed by Kambara and Takahashi at Hitachi [2], electrophoresis tracks consist of upper and lower capillary arrays separated by the sheath buffer solution. The upper capillaries are 100 μm ID and 200 μm OD and the lower capillaries are 200 μm ID and 300 μm OD. The upper and lower capillaries are 1 mm

apart and are aligned to coincide. The sheath flow solution, which draws each sample stream into the corresponding lower capillary tube, forms well-isolated and equally spaced sample streams.

However, a simpler expedient is to construct a tapered sheath flow cuvette. The narrow walls of the cuvette is designed to be 50 μm wider at the top than at the bottom (Fig. 4.1). As the linear array of capillaries is inserted into the cuvette, they are squeezed together and equally spaced. This compact design should allow more capillaries to be arrayed compared with the one from Hitachi.

A subtle yet important aspect to the well defined sample streams is the quality of capillary ends that are inserted into the sheath flow cuvette. Equally spaced sample streams are usually obtained when the detection ends of the capillaries are well trimmed. Capillary cut with a micro-dicing saw (which is a sophisticated version of a glass cutting saw used in the semiconductor manufacturing process to cut apart chips from a silicon wafer) gives a very flat end; however, after some practice, we can manually trim the capillaries to make fairly flat ends.

4.2.2. LASER BEAM PROFILE

In this multiple capillary system, a single laser beam is aligned to be parallel to the long axis of the cuvette, simultaneously exciting fluorescence from each sample stream in the cuvette. The size of the laser beam should be chosen to ensure similar illumination of each sample stream.

Wavelength (λ), beam waist spot-size (ω_0) and the location of the beam waist are characteristic of a Gaussian beam. Simple optical elements can transform the Gaussian beam into a new Gaussian beam with a newly defined beam waist and location. The transformation follows the ABCD law which assigns a 2×2 matrix

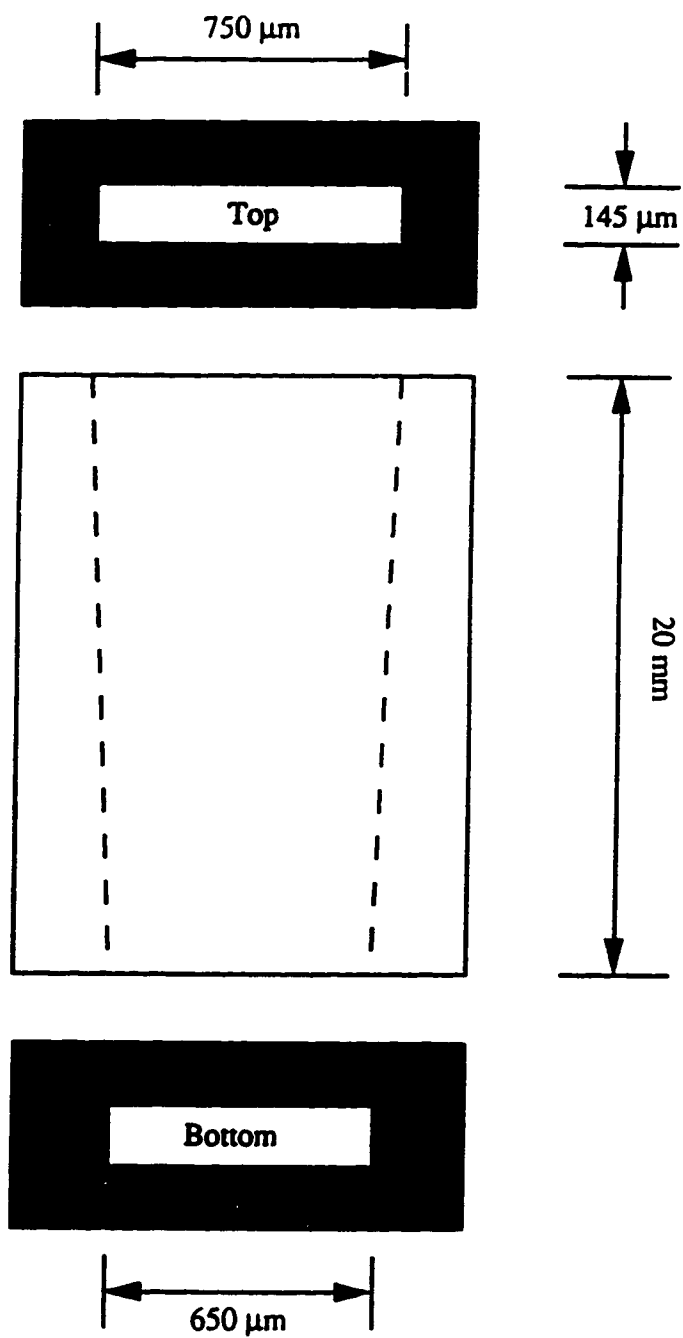


Fig. 4.1. A tapered sheath flow cuvette for an array of five capillaries, nominally $50\ \mu\text{m}$ ID and $150\ \mu\text{m}$ OD.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad (4-1)$$

to the specific optical element. For example, the matrix for a straight section with length of d would be

$$\begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \quad (4-2)$$

and the matrix for a thin lens with focal length of f would be

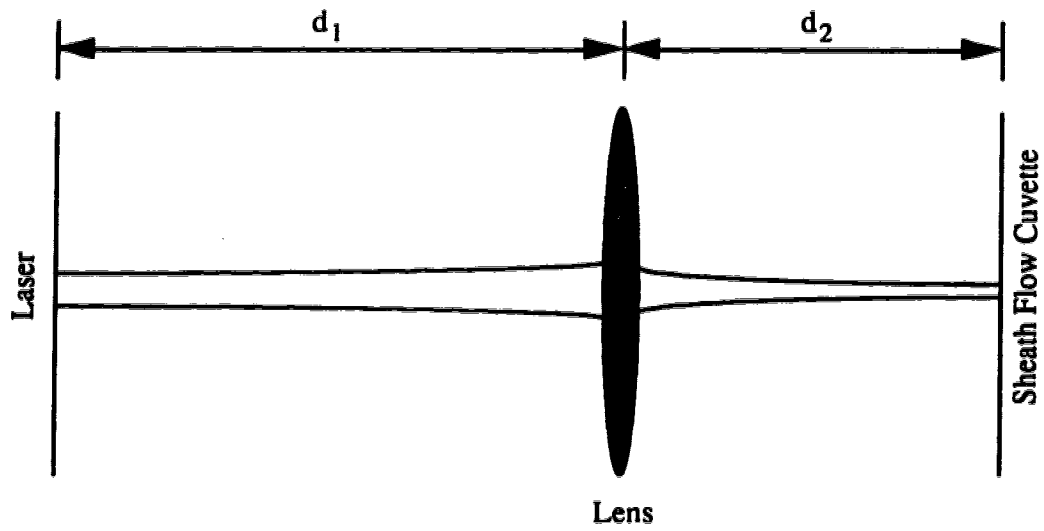
$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \quad (4-3)$$

The beam spot size, ω , at any point along the beam path is given by

$$\omega = \omega_0 \sqrt{A^2 + \left(\frac{B}{Z_c}\right)^2} \quad (4-4)$$

Z_c is called the confocal distance, $Z_c = \pi n \omega_0^2 / \lambda$, where ω_0 is the spot size at the beam waist, n is the refractive index of the medium, and λ is the wavelength of the light.

This law may be applied to a simple system where a laser beam is focused onto the sheath flow cuvette by a lens with a focal length, f , placed a distance of d_1 away from the laser output coupler (Fig. 4.2). If, for simplicity, the refractive indices of the wall of the sheath flow cuvette and the aqueous solution are not taken into account, the beam spot size should be



$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \times \begin{bmatrix} 1 & d_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 \times d_2}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{bmatrix}$$

Fig. 4.2. A schematic of the transformation of a Gaussian beam by a thin lens. Each element is assigned a 2×2 matrix and all of them combined to form a new matrix.

$$\omega = \omega_0 \sqrt{\left(1 - \frac{d_2}{f}\right)^2 + \frac{\left(d_1 + d_2 - \frac{d_1 \times d_2}{f}\right)^2}{Z_c^2}} \quad (4-5)$$

Equation 4-5 can be differentiated with respect to d_2 and set equal to zero to give the transformed spot size at beam waist, ω_0' , which is given as

$$\omega_0' = \frac{\omega_0}{\sqrt{\left(\frac{Z_c}{f}\right)^2 + \left(1 - \frac{d_1}{f}\right)^2}} \quad (4-6)$$

where ω_0 is the untransformed laser beam waist. When the input laser beam waist is known, the transformed beam waist, ω_0' , is a function of f , the focal length of the focusing lens, and d_1 , the distance between the laser output coupler and the focusing lens.

The spot-size, ω , of the transformed laser beam a distance Z from the location of the beam waist ($Z = 0$) is given by

$$\omega' = \omega_0' \sqrt{1 + \left(\frac{Z}{Z_c}\right)^2} \quad (4-7)$$

This equation indicates that the beam spot size will expand by 40% at a distance Z_c from the beam waist.

Cannon *et al.* [3] and Cohen *et al.* [4] have described several techniques for the determination of spot size of a Gaussian beam. Edges, wires, slits or rulings were the objects used to scan across the beam. The change in beam intensity was detected, and the spot size at beam waist was calculated.

In this experiment, the intensity of the argon ion laser beam was measured with a power meter (Lexel, Model 504). A razor blade was mounted on a two-axes translation

stage and the razor edge was scanned across the focused laser beam with 5- μm steps. The ratio of the transmitted power P , to the total power, P_0 is

$$\frac{P}{P_0} = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x}{\sqrt{2}\omega} \right) \right] \quad (4-8)$$

where ω is the beam radius (defined as the distance between the beam center and the point where the irradiance falls to $e^{-2} = 13.5\%$ of the maximum irradiance) and erf is the error function. This process was repeated at different positions along the laser beam.

It turned out that a $1\times$ microscope objective ($f = 73.5\text{ mm}$) would transform the laser beam into a new beam that was best suited for the multiple capillary system when the laser was 1.2 m ($d_1 = 1.2\text{ m}$) away from the focusing lens (Fig. 4.3): the beam waist had a spot size 27 μm in diameter; 0.5 mm away from the waist, the spot size was less than 30 μm . Other objectives used to focus the beam, for example, $2.5\times$ or $4\times$, gave smaller spot size at the beam waist and larger divergence, making uniform illumination of multiple sample streams difficult.

It should be noted that a $1\times$ microscope objective does focus the well collimated laser beam and the location of the transformed beam waist is not necessarily one focal length from the lens. It should also be noted that this focused laser beam is useful even for an array of 30 capillaries with 150- μm OD because over a distance of 3 mm, the beam spot size is still within 50 μm .

4.2.3. FLUORESCENCE DETECTION OPTICAL TRAIN

The fluorescence detection optical train in this multiple capillary system consists of collecting optics, a filter (or filter wheel) and an array of photodetectors. This instrument design uses a well-corrected, flat-view microscope objective as the collection optic. The detection sensitivity scales up with fluorescence collecting efficiency of the

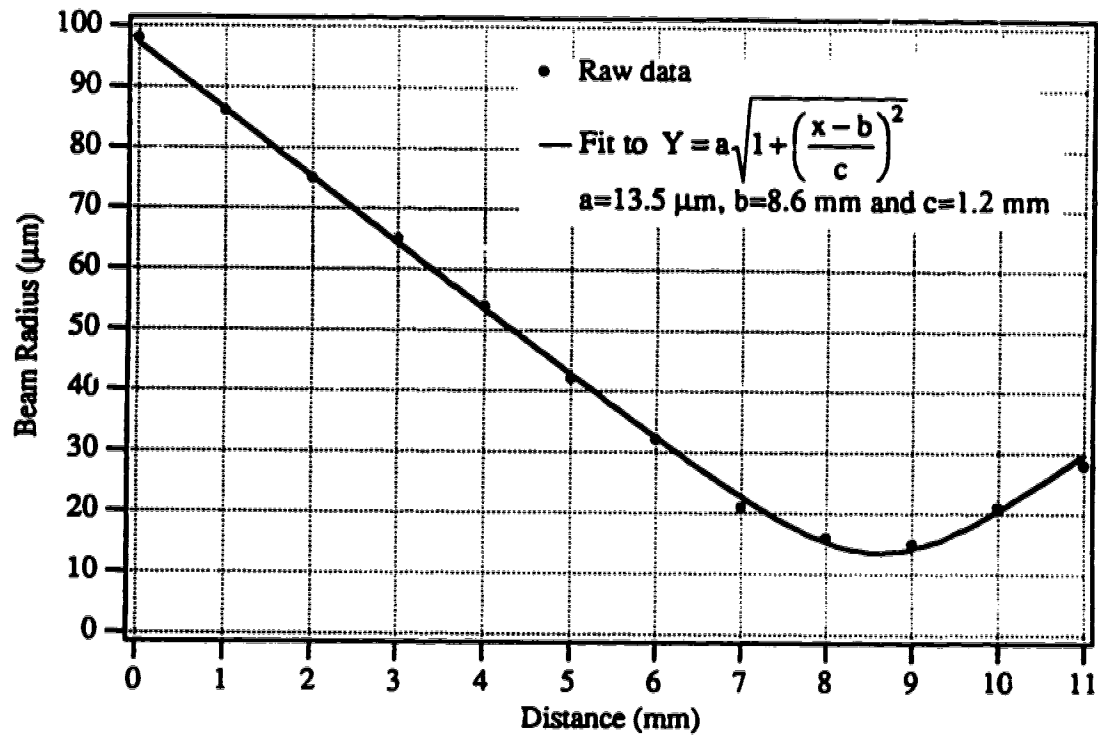


Fig. 4.3. The laser beam profile of the optical system shown in Fig. 4.2. The laser is Model 2211-20SL (Uniphase) and the focusing objective is 1×, $f = 73.7 \text{ mm}$ (Melles Griot).

microscope objective. Fluorescence collection efficiency is related to the numerical aperture (NA) of the collecting microscope objective:

$$\text{collection efficiency} = \sin^2 \left[\frac{\arcsin(\text{NA} / n)}{2} \right] \quad (4-9)$$

where n is the refractive index of the surrounding medium. Usually the objective is surrounded by air and $n = 1$. Table 1 lists the collection efficiency for lenses of different numerical apertures.

As we can see, high numerical aperture lenses generate high collection efficiencies. However, there are two constraints related to the high numerical aperture lenses. First, few lenses are designed with both high numerical aperture and long working distances. Expensive objectives with reasonable numerical aperture are available with a working distance greater than 3 mm such as the one used in the instrument described in Chapter 2. Second, high numerical aperture lenses usually come with high magnification. In the multiple capillary system, a wider field of view would require use of lower power objectives. For these reasons, an inexpensive microscope objective, $20\times$, 0.5 NA with a working distance of 1.44 mm, is the choice in this instrument. Objectives with working distance as long as 8 mm and NA of 0.4 may be useful for even larger array of capillaries.

Because the sheath flow cuvette window is only 1-mm thick, a $20\times$, 0.50 numerical aperture microscope objective, which has a working distance of 1.44 mm, can

Table 4.1

Collection Efficiencies of Different Lenses ($n = 1$)

NA	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Collection efficiency	0.50	0.28	0.20	0.14	0.10	0.067	0.042	0.023	0.010	0.003

be used to collect fluorescence from the sample streams. The 50- μm ID capillaries have 150- μm OD; the image produced by the collection optic is an array of 1-mm diameter spots, spaced on 3-mm centers.

An array of five discrete SiAPDs were used as photodetectors. The individual SiAPDs provide several advantages over the use of a photodiode array or a CCD camera. The SiAPDs have extremely high quantum efficiency plus internal gain. The discrete devices are individually amplified. A multiplexed A/D converter can record the signal from many SiAPDs at very high rates, which is useful for high speed separations of sequencing fragments.

These SiPADs, configured as photon counting modules by EG&G, had very low dark counts (Fig. 4.4A), and very high quantum efficiencies (up to 50% at 500 nm). These devices have the capability of single photon detection. However, the background signals for these photodetectors were 2 to 3 orders higher than dark counts (Fig. 4.4B). This is mainly due to the Raman scattering over this wavelength range. The detection sensitivity will be ultimately limited by the Raman scattering in the aqueous solution unless some other technique is used such as pulsed laser excitation and time-gating detection.

However, the choice of discrete SiAPDs leads to an experimental challenge. The SiAPDs, which have a 0.5-mm useful diameter or 0.2-mm² round active area, are provided in 6-mm wide cans. This aspect ratio is not compatible with the 50- μm width and 150- μm spacing of sample streams in the sheath flow cuvette; fluorescence can not be imaged directly onto the photodiodes. Instead, an array of optical fibers couples fluorescence into the APDs. However, the optical fibers have 100- μm cores, which is much smaller than the 1-mm diameter image of the sample stream produced by the 20 \times microscope objective. Gradient index (GRIN) lenses, with 1- to 2-mm diameter and 0.25 pitch, very efficiently couple the 1-mm fluorescence images into the fiber optics. Because of these requirements, each photodetector consists of a GRIN lens, a fiber optic

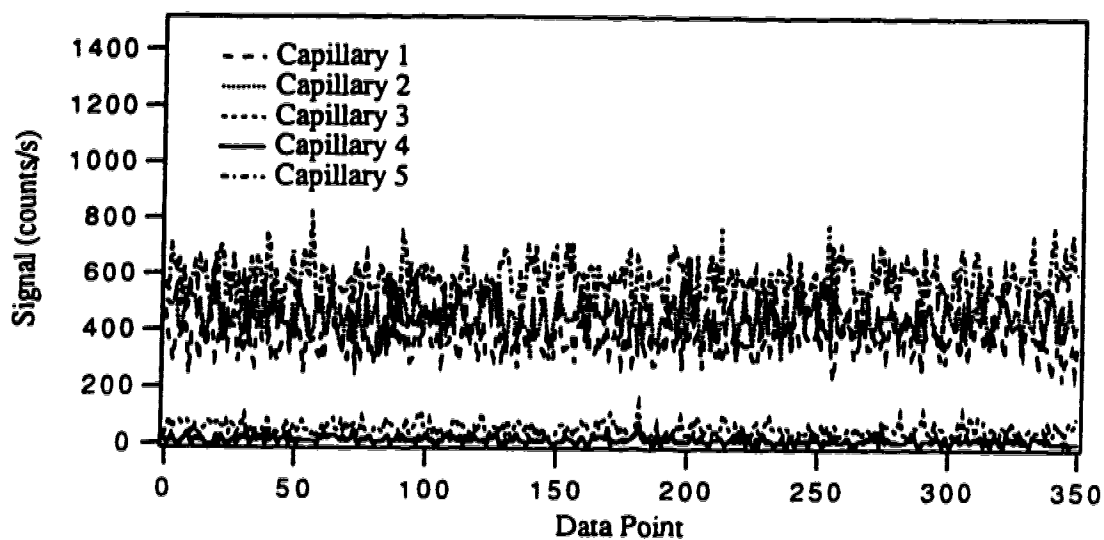
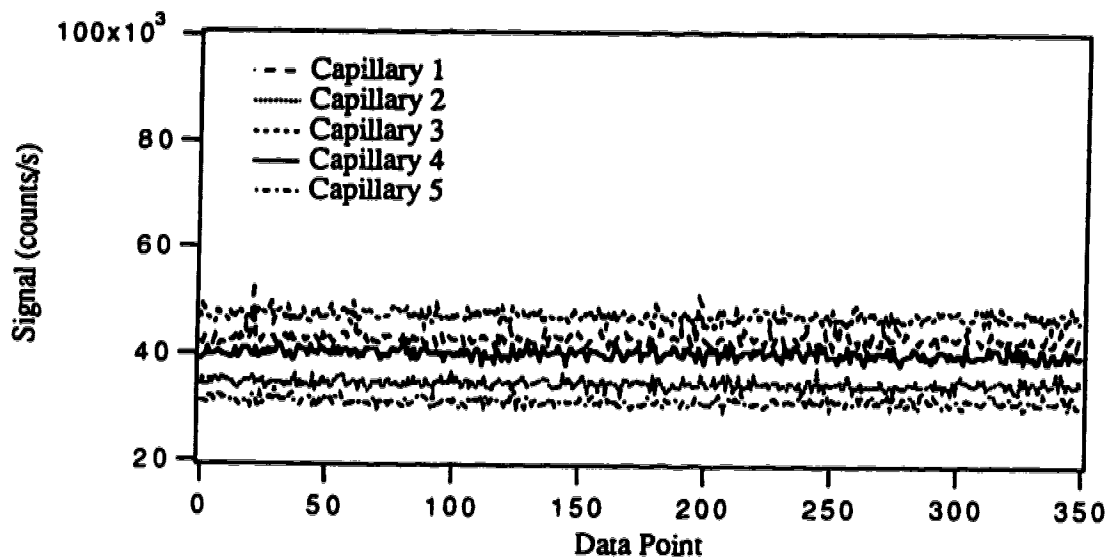
A. Dark Counts**B. Background Signal**

Fig. 4.4. Dark counts and background counts of the photon counting modules. (A) Dark counts, and (B) background counts, i.e., the signals with the laser on.

and a SiAPD; five of this combination were used in the five capillary system.

A technical difficulty came from the space constraint. An array of GRIN lenses was aligned, 3-mm apart, on the image plane of the microscope objective, yet these GRIN lenses must be concentrically aligned with the corresponding optical fibers which had 100- μm core and 140- μm cladding. Special connectors were needed to couple the GRIN lenses and optical fibers together. The connectors used in the previous chapter were home-made and they were not useful for high sensitivity fluorescence detection. Fortunately, we obtained a connector commercially from NSG (Collimator, M-type). After a few modifications, the GRIN lenses efficiently coupled fluorescence into the optical fibers and the 3-mm spacing was no longer a problem.

This commercial collimator came with fiber sleeve, GRIN lens holder and GRIN lens that was 1.8-mm diameter and 0.25 pitch. The GRIN lens had already been fixed into the holder. The optical fibers were treated with a fiber cleaver (Model F-BK2, T&B, SC, USA), producing perpendicular, flat end-faces. The fibers were inserted into the sleeves, glued with epoxy and then fitted into the GRIN lens holders to form single units. All these operations were carried out under a microscope.

The array of GRIN lenses were placed at the image plane of the collection objective not in a parallel fashion, but instead, in a fan shape (Fig. 4.5). As we can recall from chapter 3, a GRIN lens has the ability to transform a well collimated beam into a point spot or visa versa, all GRIN lenses should point to the center of the collecting microscope objective for better fluorescence collection.

The use of fiber optics has proven to be valuable in alignment of the system. The fiber optics may be disconnected from the SiAPDs and illuminated with a lamp, creating a perfect back-illumination of the optical system. When viewed through an alignment microscope placed opposite the sheath flow cuvette from the collection optic, the illuminated fiber optics transmit light through the GRIN lenses to the 20 \times microscope objective into the sheath flow cuvette. Alignment was achieved by flowing dilute

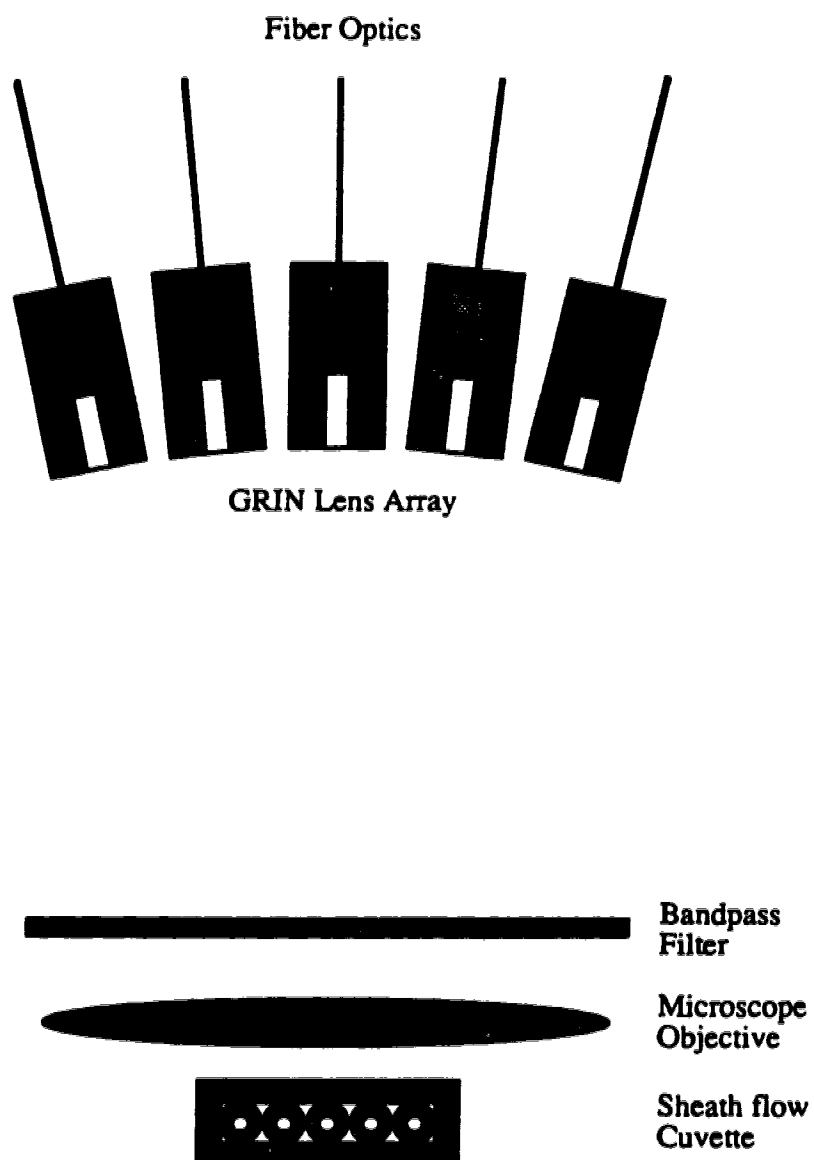


Fig. 4.5. The arrangement of GRIN lens array in the image plane. The GRIN lenses were arranged in a fan shape in order to increase the fluorescence collection.

fluorescent dye through the capillaries; the illuminated spots from the sample and the back-illuminated spots from the fiber optics were superimposed (Fig. 4.6). The fiber optics were then re-connected to the SiAPDs, and a final tweaking of the optical system was performed.

4.3. EXPERIMENTAL

4.3.1. MATERIAL

Capillaries, 50- μm ID and 150- μm OD from Polymicro Technologies (Phoenix, AZ, USA) were used as received with no pretreatment. Acrylamide and N,N,N',N'-tetramethylethylenediamine (TEMED) were electrophoresis-purity reagents from Bio-Rad (Mississauga, Canada). Tris base and boric acid were ultrapure reagents (Sigma, St. Louis, MO, USA). Urea was enzyme grade (Life Technologies, MD, USA). Borax and EDTA were analytical-reagent grade (BDH, Edmonton, Canada) and ammonium persulfate was ultrapure electrophoresis grade (Boehringer Mannheim, Laval, Canada).

Recently, ABI has introduced a new set of dye-labeled dideoxynucleotides [5]. These so-called "PRISM" chain terminators are designed for use with T7 DNA polymerase. The sequencing reaction was carried out in 40 mM MOPS buffer, pH 7.5, 50 mM NaCl, 10 mM MnCl_2 , and 15 mM sodium isocitrate. 0.8 pmole of primer (Applied Biosystem-21M13) was annealed to 2 μg of single-stranded pGEM-3ZF template at 65 °C for 2 min. followed by slow cooling to room temperature. A mixture of dye-dideoxy- and deoxynucleoside triphosphates (T7 terminator mix) was added to give an average nucleoside ratio (dNTP/ddNTP) of greater than 300:1 with dNTP-S used in place of dNTP. After the mixture was warmed to 37 °C, 1.5 units of DNA T7 polymerase and 0.006 units of pyrophosphatase were added. Incubation continued at 37 °C for 10 min, after which the DNA was precipitated with ethanol. The samples were ethanol-washed

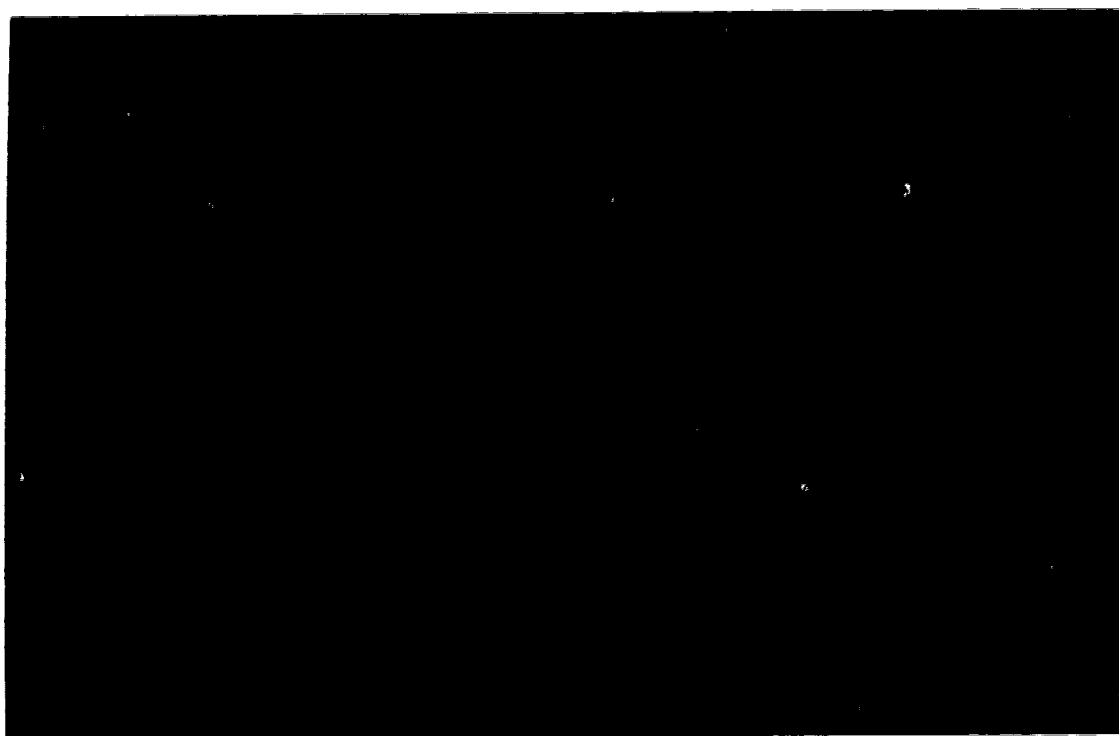


Fig. 4.6. (a) The image of the GRIN lens array. The image was the back-illumination of optical system seen at the opposite side of the sheath flow cuvette through an auxiliary microscope.

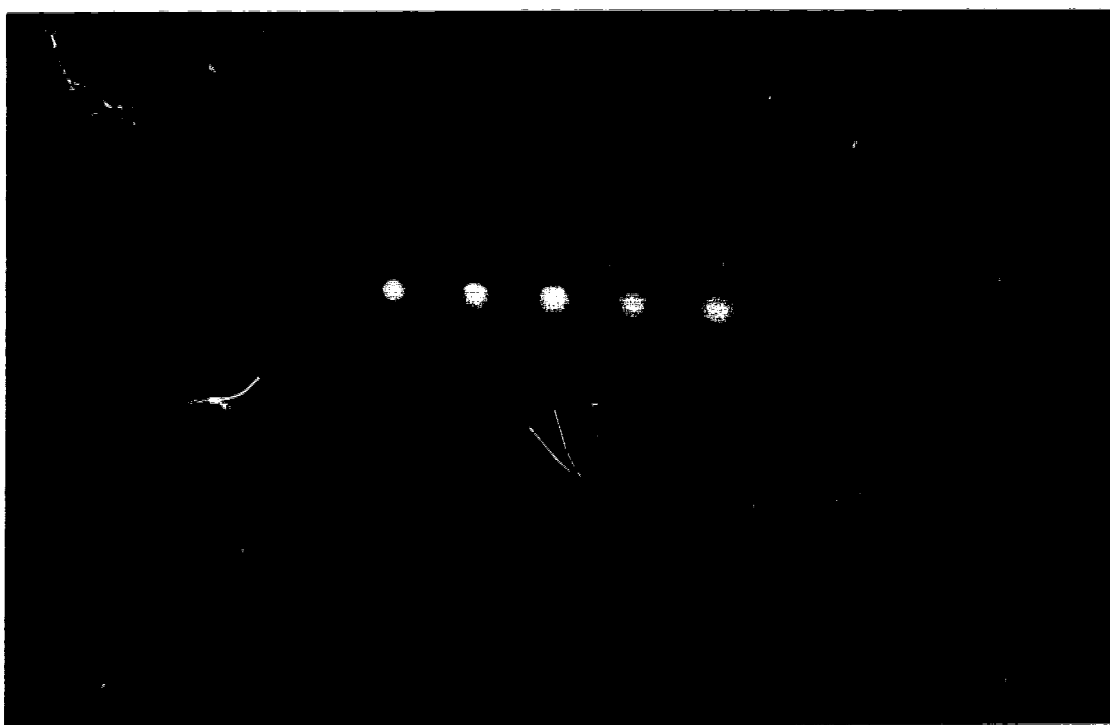


Fig. 4.6. (b) The image of superimposition of the fluorescence spots and back-illuminated spots as seen through an auxiliary microscope placed opposite the sheath flow cuvette from the collecting optic.

and resuspended in a 4- μ l 49:1 mixture of formamide:0.5 M aqueous EDTA, heated at 95 °C for 2 minutes, and injected onto the capillaries by applying a 100 V/cm electric field for 30 seconds.

4.3.2. METHODS

A stock borate solution was prepared by dissolving 0.478 g of borax in 50 ml water. It was diluted to a final concentration of 10 mM. A stock 10 \times TBE (pH 8.3) solution was prepared by dissolving 108 g Tris, 55 g boric acid and 40 ml of 0.5 M EDTA in water to a final volume of 1 liter. By using a gel-filling apparatus, capillaries, 38 cm long, were filled with degassed 4% - 8% acrylamide, 8 M urea, 0.07% (w/v) ammonium persulfate, 0.07% (v/v) TEMED, solution prepared in a TBE buffer (1 \times TBE final). Polymerization reaction was allowed to proceed overnight.

Electrophoresis operation was carried out at room temperature. A 30-kV power supply (Spellman, Plainview, NY, USA) was used to drive electrophoresis. There was a polarity reverse of the applied voltage when separation mode changed from free zone electrophoresis to gel electrophoresis. 10 mM borate (for free zone electrophoresis) and 1 \times TBE (for gel electrophoresis) were used as the sheath fluids. The sheath flow was provided by a precision syringe pump (Model 906, Harvard Apparatus, MA, USA), or simply by siphoning with the height difference of 7 to 2 cm between sheath flow reservoir and waste container. Sheath flow rate was typically 0.3 ml/h.

4.4 RESULTS AND DISCUSSION

4.4.1. INSTRUMENT PERFORMANCE

Given the above challenges in the design process, I was very gratified with the

performance of the multiple capillary system. The detection limits for the five-capillary detector were evaluated with fluorescein in free zone electrophoresis. The 50- μm ID, 150- μm OD capillaries were filled with a 10-mM borate buffer, pH 9.2. Separation was at 300 V/cm across the 37.0 cm long capillaries; 1.1 nl sample were injected electrokinetically (1.0 kV, 5 sec). A series of four concentrations of fluorescein were prepared, from 2×10^{-11} to 2×10^{-12} M. Fluorescence was excited by a 4.0-mW argon ion laser beam at 488 nm, collected with a 20 \times , 0.5 NA objective. Fig. 4.7 presents an electropherogram of a 2×10^{-12} M solution of fluorescein (2.2 zeptomoles or 1,300 molecule injected). The five traces were recorded simultaneously. The difference in migration time reflects the difference in the nature of the capillaries as discussed in the previous chapter. The peak area corresponds to 12,000 photons above the background signal level; each molecule generates an average of 9 detected photons. Blank injections were performed by dipping the capillaries into the dye solution without application of injection potential; no peaks were detected from these blanks. Detection limits (3σ) were 130 ± 30 molecules (2×10^{-13} M) injected onto the capillaries.

This detection limit data is among the best that has been published in fluorescence detection in capillary electrophoresis, even compared to the data generated with single capillary systems [6]. This performance reflects the high sensitivity of the detection system. First, the detection system generated a low background signal level with the SiAPDs operated at 0 °C. Background levels ranged from 5,000 to 10,000 counts in a period of 0.2 s with a 4.0 mW argon ion laser as the source. The shot noise ranged from 70 to 100 counts. Second, the optical system had a high photon detection efficiency. The 20 \times , 0.5 NA microscope objective had a collection efficiency of 6.7%, which is about half of the collection efficiency of the 32 \times , 0.65 NA objective used in the single capillary instrument. This lower collection efficiency was well compensated by the high quantum efficiency of the SiAPD, 45%, as compared to 15% for a PMT at the fluorescence wavelength.

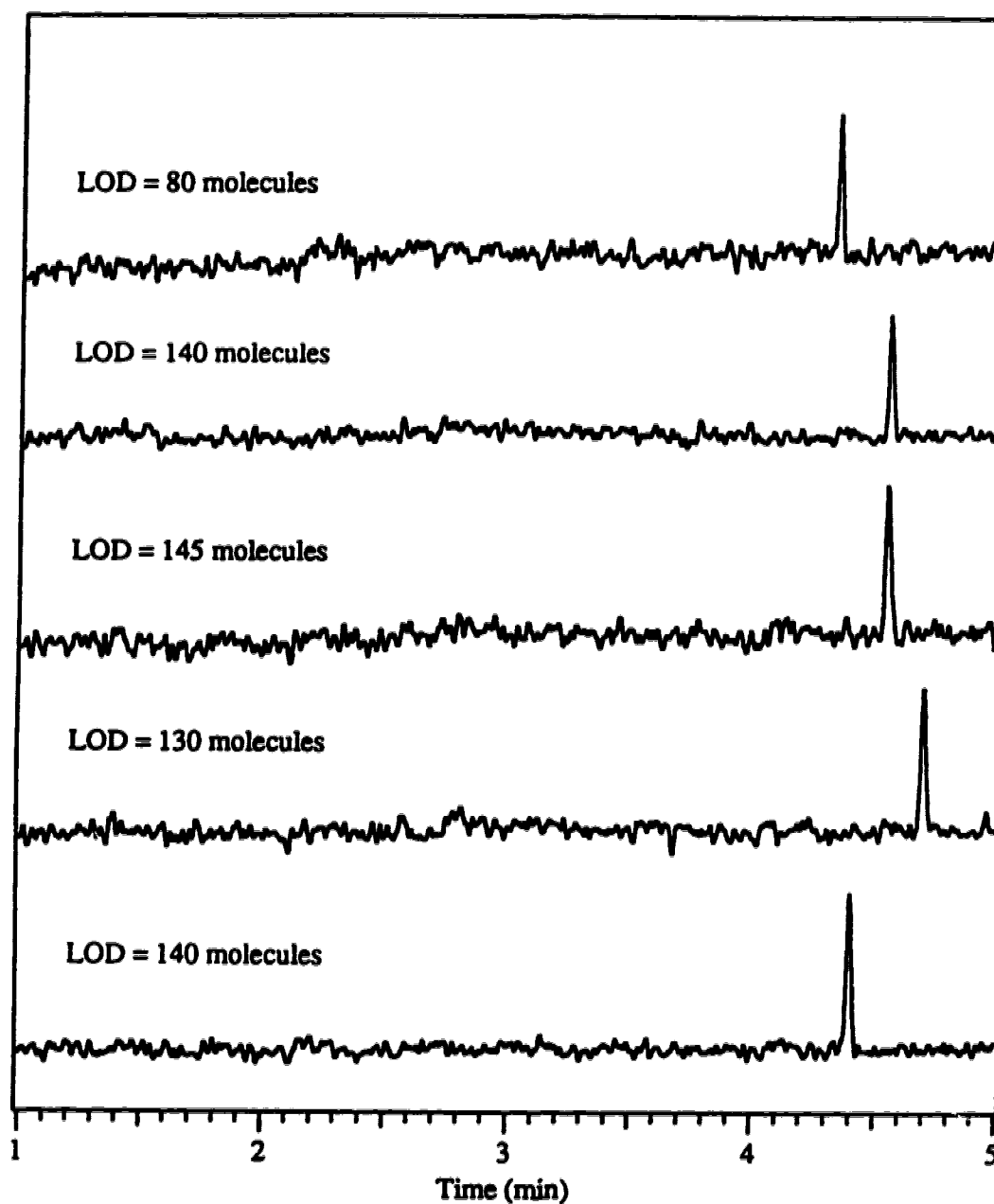


Fig. 4.7. Injection of 1300 fluorescein molecules. The capillaries were 50 μm ID, 150 μm OD and 37.5 cm long. Fluorescein, 2×10^{-12} M, was injected at 1 kV for 5 seconds. The electrophoresis was conducted at an electric field strength of 300 V/cm. The laser power was 4.0 mW. Each data point was a 0.2-second count. The data were subjected to an digital filtering.

In the detection system, the $20\times$, 0.5 NA microscope objective had a collection efficiency of 6.7%; over the whole fluorescence spectrum, the bandpass filter had a transmission of about 40%, and the SiAPD had an average quantum efficiency of 45%. Detection of 9 photons from each fluorescein molecule implies each fluorescein emitted 750 photons in passing through the laser beam.

It should be pointed out that the system should provide even better detection limit if the excitation and detection are shifted to longer wavelengths. In general, lower Raman background is observed for longer-wavelength excitation because the Raman intensity scales with λ^{-4} . Meanwhile, SiAPD quantum efficiency is also higher at higher wavelength (see Table 3.1 in Chapter 3). For example, use of the green He-Ne laser ($\lambda = 543.5$ nm) for excitation of rhodamine 6G (fluorescence emission maximum at 575 nm), or the yellow He-Ne laser ($\lambda = 594$ nm) for sulforhodamine 101 (fluorescence emission maximum at 607 nm) should provide superior results. In fact, Li *et al.* has reported single molecule detection by using a SiAPD as a detector [7]. By using SiAPD, 80 photons can be detected from single rhodamine 6G molecule instead of 10 photons detected by a PMT.

The peaks generated in the five capillaries have Gaussian shape and are fitted with the following equation

$$Y = K_0 + K_1 \exp[-(t - K_2)^2 / K_3^2] \quad (4-9)$$

where K_0 , K_1 , K_2 , K_3 are constants. K_0 is the background level, K_1 is the peak height, K_2 is the migration time of the peak and K_3 is the standard deviation of the peak. Y represents the fluorescence intensity, and t is time. The separation efficiency in terms of theoretical plates, N , is given as

$$N = 2 \left(\frac{K_2}{K_3} \right)^2 \quad (4-10)$$

Calculation shows that N ranges from 690,000 to 1,000,000 per meter for a typical run (Fig. 4.8) when fluorescein migrated through 5 capillaries, 37.5 cm long and filled with 10 mM borate buffer (pH9.2), under an electric field strength of 300 V/cm.

The use of discrete capillaries and photodetectors results in some compromises in the system:

First, there is a compromise to achieve similar signal levels. It should be noted that even though the source was the same, there were differences in signal levels as can be seen in Fig. 4.4. Alignment was performed to generate similar sensitivity for each detection channel.

Second, there is a compromise between laser power and linear range. The count rates of these photodetectors can reach the maximum of 1 to 2 megacounts per second. Beyond this level, the rate curve begins to bend due to the dead time effect. As the background level is on the order of 10^4 counts per second, the practical analysis range is only about two orders of magnitude. Although signal-to-noise ratio increases modestly with higher laser power (Fig. 4. 9), a compromise must be made between linear range and the laser power. For this reason, 4.0 mW, the lowest power for the laser, was generally used in these experiments. Good electropherograms for DNA sequencing can be obtained after the compromise. To extend the linear range, some other low power lasers, such as green or yellow He-Ne lasers can be used, of course, for some other dyes.

Third, there is a compromise between counting interval and the counter's maximum counts. To some extent, signal-to-noise ratio also scales with longer counting time (or integration time) under shot noise limited condition. However, there is another compromise between counting interval and counter's maximum reading, which is 16 bit or 65536 for those on NB-TIO-10 board; 0.1 to 0.2 second is used for general analysis

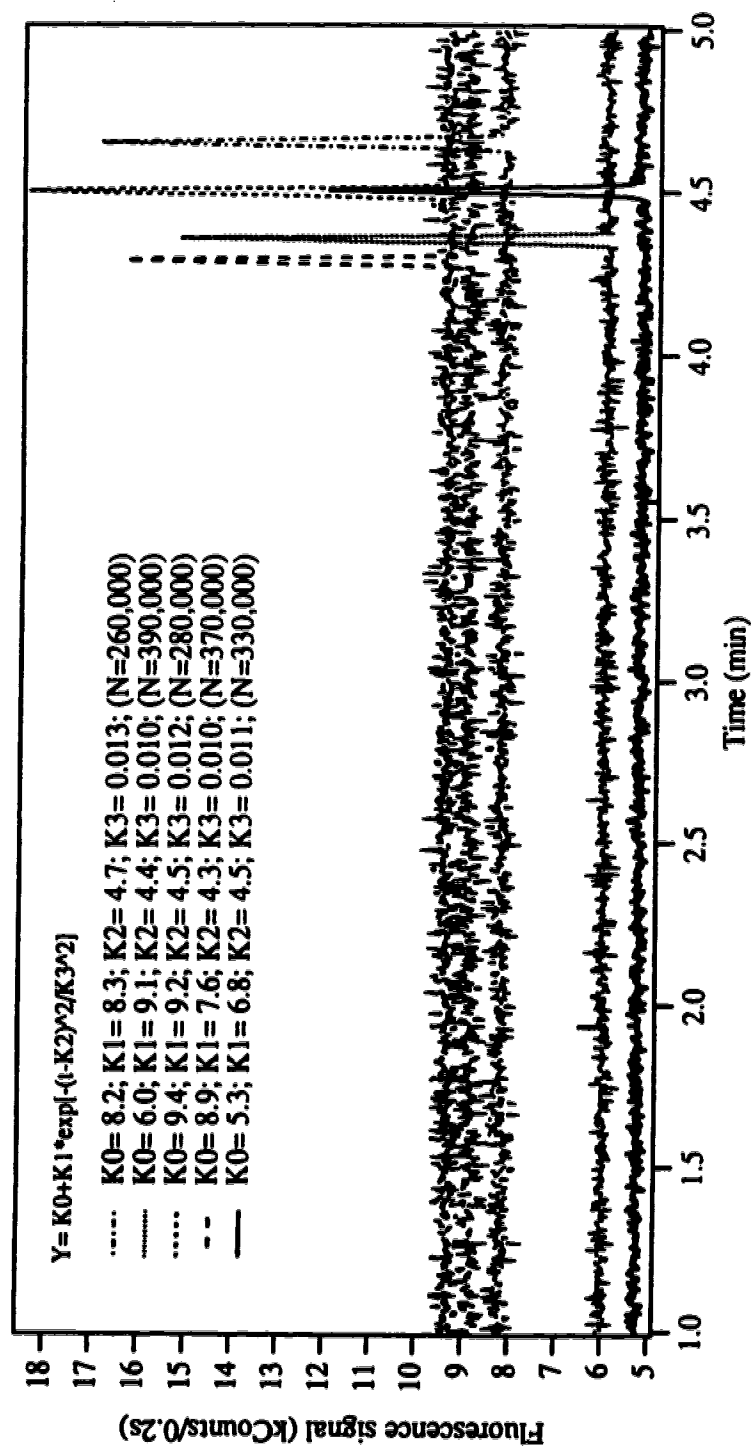


Fig. 4.8. Electropherogram of fluorescein. The capillaries were 50 μm ID, 150 μm OD and 37.5 cm long. Fluorescein, 2×10^{-11} M, was injected at 1 kV for 5 seconds. The electrophoresis was conducted at an electric field strength of 300 V/cm. The laser power was 4.0 mW. Each data point was a 0.2-second count. The results of regression analysis of Equation 4-9 to the data is shown in the insert.

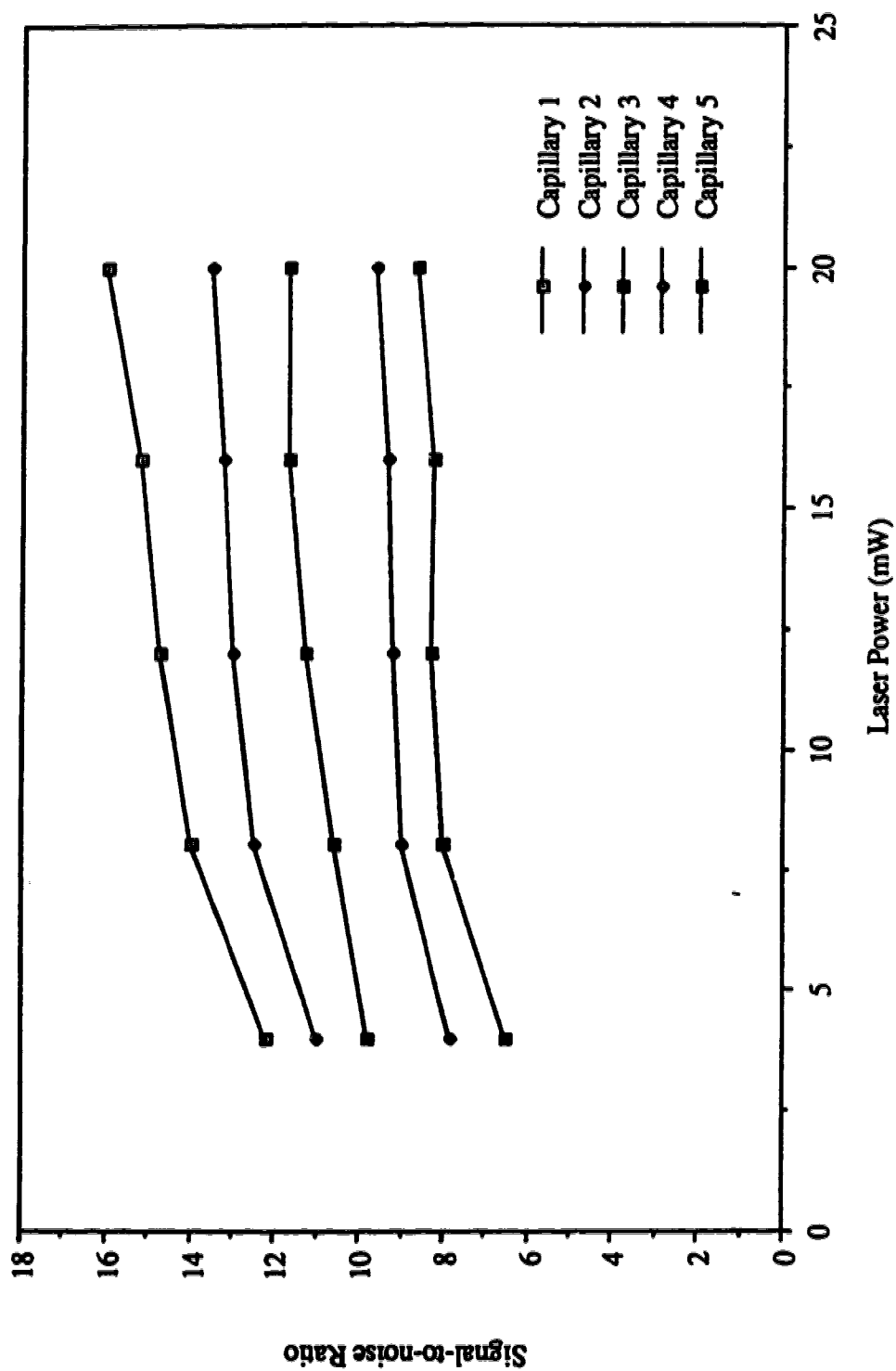


Fig. 4.9. The signal-to-noise ratio as a function of the laser power. The measurement was made with a continuous flow of 2×10^{-11} M fluorescein in borate buffer, pH 9.2, at an electric field strength of 300 V/cm.

such as single base sequencing. The count rate can be increased to perform DNA sequencing by capillary gel electrophoresis at the highest speed to date, a speed of 53 bases per minute [8].

4.4.2. GEL

In gel electrophoresis, the gel provides a sieving environment for size-dependent separations. For separation of DNA sequencing fragments in slab gel electrophoresis, the gel matrix is almost exclusively cross-linked polyacrylamide (although some proprietary gels such as Long Ranger and Hydrolink are available). The quality of the gel in capillary format is an important issue. Poppe and co-workers described a photopolymerization method for preparation of gel-filled capillaries [9]. Novotny and co-workers reported a progressive polymerization method, in which ammonium persulfate and TEMED were electrophoretically introduced into the capillary [10]. Karger's group introduced a vacuum method to prepare gel-filled capillary [11]. Baba reported a gel prepared in the same way [12]. Mathies' group used the vacuum method to produce gel-filled capillary array [13].

To succeed in DNA fragment separation, care must be taken in gel preparation and sample separation. First, inside the capillary, the gel experiences an electroosmotic pressure, leading to the extrusion of the gel out of the capillary from the injection end [14]. Sometimes, the gel also experiences electrostriction, which results in the migration of gel out of the detection end of the capillary [14], blocking the laser beam in the sheath flow chamber. These undesirable movements can be eliminated through covalent bonding of the gel to the capillary wall, either for the whole length or for one end of the capillary. Second, degassing the gel is necessary, because oxygen tends to quench the free-radical polymerization. Third, DNA sequencing fragment separation will be the ultimate test of the gel quality. Sometimes a bubble-free gel, checked under the microscope, doesn't provide desirable results. After injection of sample and application of high voltage, there

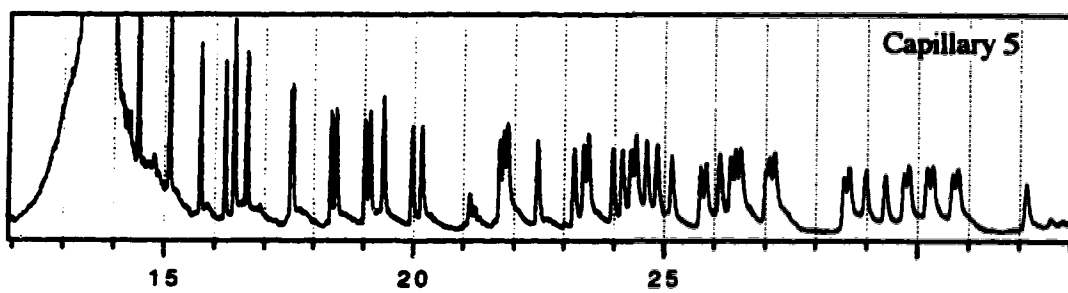
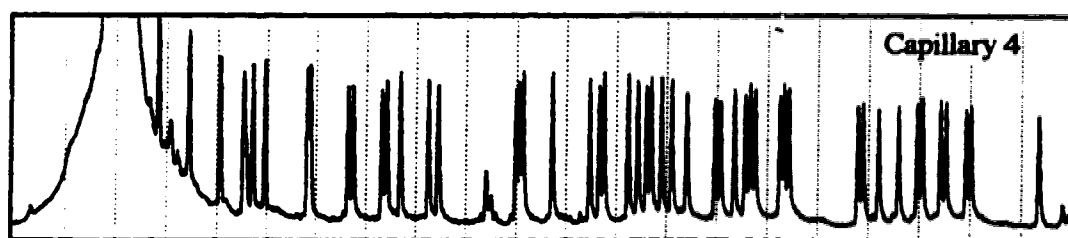
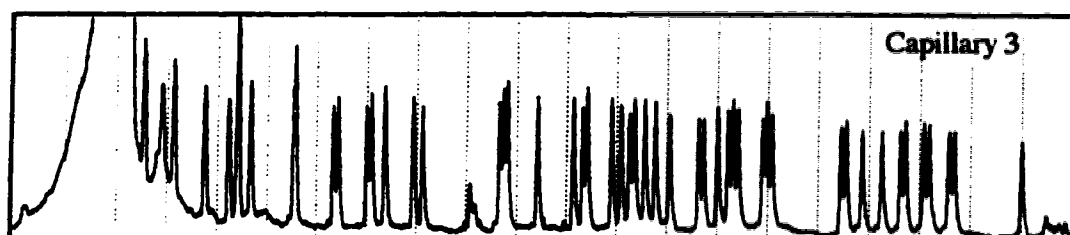
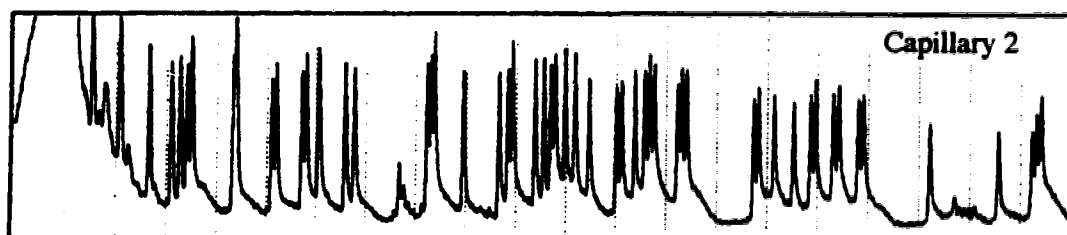
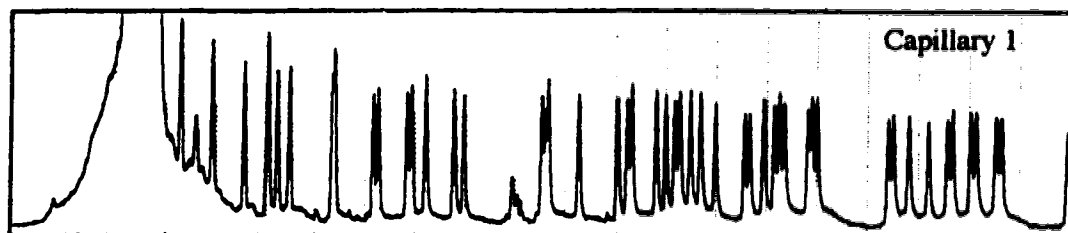
is a change in environment at the injection end of the capillary, such as temperature, ionic strength and conductivity, which degrades the separation. Besides, large fragments such as templates tend to block the gel [15,16], and high voltage will stretch the gel. All these changes will induce stresses to the gel, leading to gel anomalies.

When gel anomalies happen, there is a sudden drop in current. For a single capillary, this current drop is reflected as a voltage change across a sampling resistor in between the sheath flow cuvette and ground. The cure is to trim a few mm of capillary down from the injection end as can be seen in Chapter 2. The high voltage is then adjusted to the capillary length and electrophoresis is resumed. For a multiple capillary system, it is difficult to monitor the current in each capillary at the high voltage end. The voltage across the sampling resistor only reflects the gross current across the capillary array. It is difficult to identify whether the current decrease is due to the breakdown of a single capillary or because of the current fluctuations from the capillary array. Even though the breakdown is identified, it is not practical to trim the specific capillary among the capillary array.

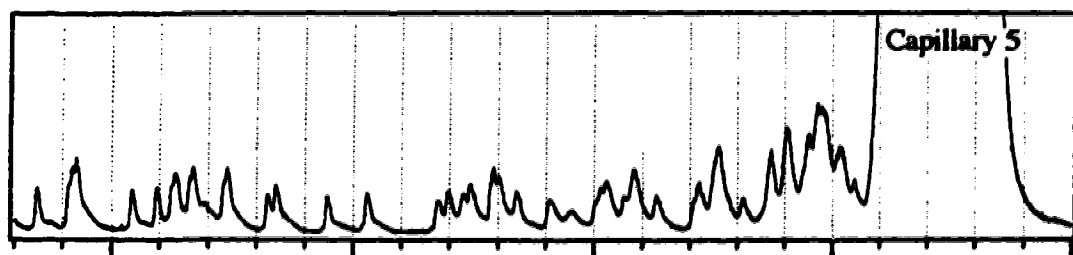
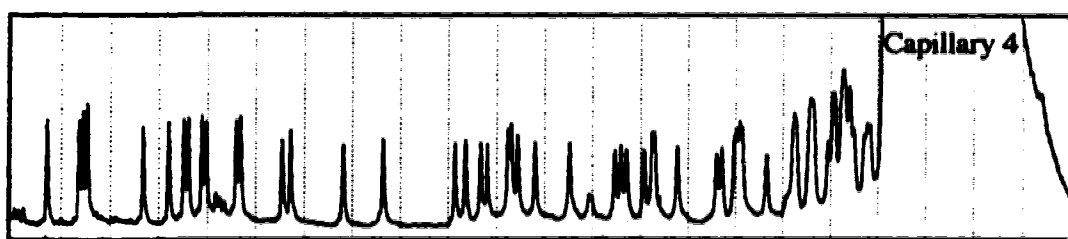
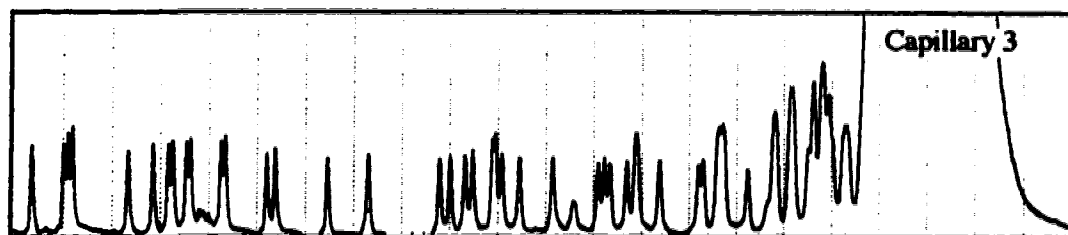
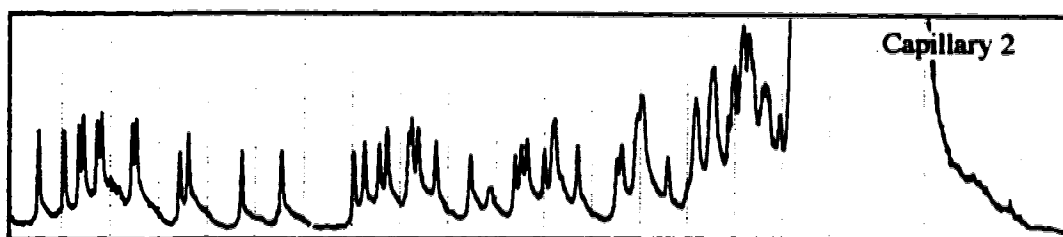
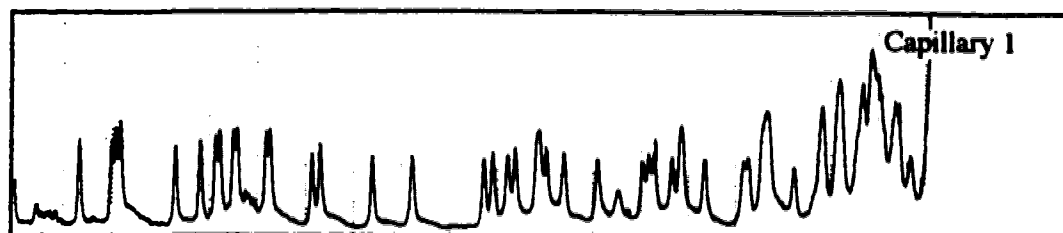
It is difficult to manage five capillary gels to achieve ideal performance when using cross-linked gels. Fig. 4.10. shows a typical separation using 5% cross-linked Long Ranger gel as the separation medium. Migration time variations ranging from 5-20% among the capillaries were observed by using cross-linked gels. This result is very similar to that reported by Kambara and Takahashi [2]. This variation in migration time arises because the cross-linked gels are rigid. When trimming multiple capillaries before and/or after sample injection, the gels may be pulled from the injection end, leaving bubbles inside the capillaries. Associated with the variation in migration times, variation in separation efficiencies can be easily observed. Separations with longer migration times usually show poorer efficiencies.

Compared to cross-linked polyacrylamide, which is rigid in nature and is poor in relieving stresses, non-cross-linked polyacrylamide, with its dynamic pore structure,

Fig. 4.10 (pages 122-123). Electropherograms of FAM-labeled, A-terminated DNA sequencing fragments generated from M13mp18. Capillaries were 50 μm ID, 150 μm OD and 33.0 cm long. The capillaries were filled with 5% Long Ranger gel. The sample was electrokinetically injected at 100 V/cm for 15 seconds and the electrophoresis was driven at an electric field strength of 300 V/cm. Capillaries were trimmed 1.0 mm from the injection end at 3 minutes of run.



Time (min)



35

40

45

50

Time (min)

functions well as a sieving medium for DNA sequencing fragment separation. It eliminates the necessity of trimming capillary after sample injection. Besides, this entangled polymer solution has relatively low viscosity and may be pumped from the capillary after use, making the capillary reusable [17]. By refilling rather than replacing the capillary, realignment of the optical system is minimized in the process of automated DNA sequencing. Crambach and Karger's groups have compared the performance of 0%C polyacrylamide at various total acrylamide concentration [18-20]. Boccek and Karger's groups reported that 0%C polyacrylamide could be pumped from a capillary [21,17]. Pentony reported DNA sequencing in 10%T, 0%C polyacrylamide at an electric field of 300 V/cm; sequence could be determined for fragments longer than 300 bases [22]. Similar results have been reported by Mathies group with 9%T, 0%C polyacrylamide [23]. Righetti and co-workers argued that the high viscosity of 10%T 0%C polyacrylamide eliminates any hope of refilling capillaries with that material; only low %T polyacrylamide has sufficiently low viscosity for replacement [24]. Guttman and co-workers reported the use of low total percent linear polyacrylamide for separation of double stranded DNA; they claimed that the capillary could be reused 100 times without replacement of the separation medium [25].

A capillary gel-filling apparatus, which is similar to the vacuum method described by Baba *et al.* [12], was built as shown in Fig. 4. 11. It enables generation of many gel-filled capillaries simultaneously. Capillaries were well-trimmed to 38 cm in length. These capillaries were first filled with silanization solution by applying vacuum. The apparatus was disconnected from the vacuum source to allow the solution to stay in the capillaries for about 20 minutes. The vacuum was reapplied to remove the silanization solution from the capillaries. Degassed linear acrylamide solution was mixed with ammonium persulfate and TEMED. One end of the capillaries was put into the acrylamide solution and the other end was put into the water as shown in Fig. 4. 11. The vacuum was applied to fill the capillaries with the acrylamide solution. Vacuum was

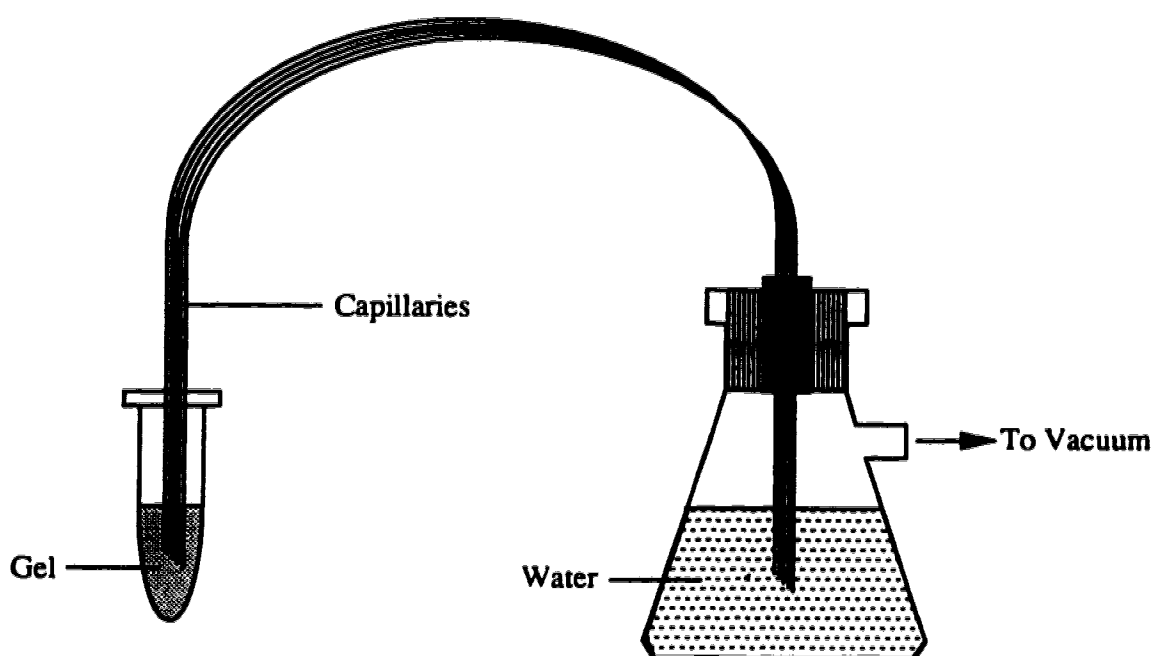


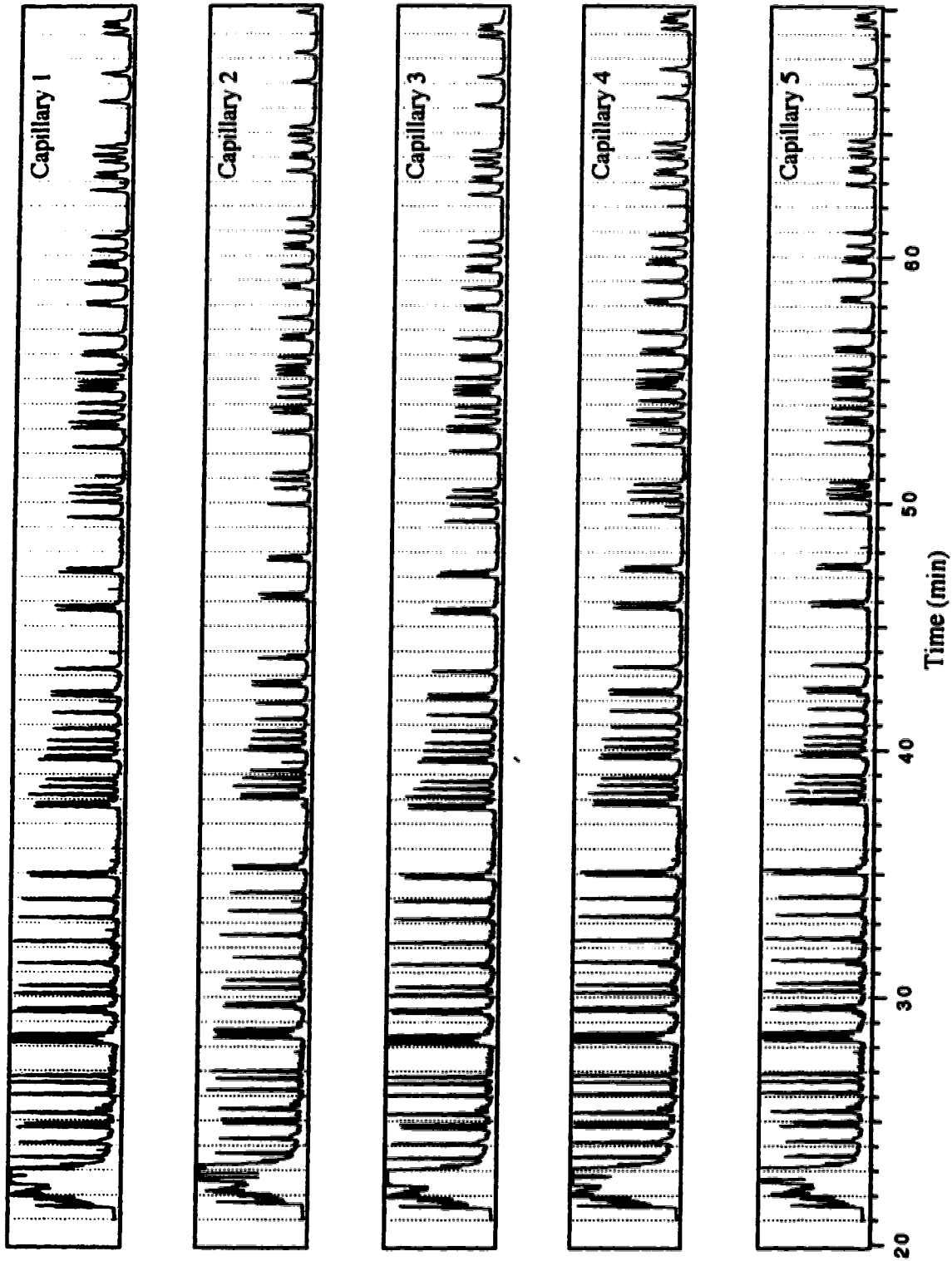
Fig. 4.11. A schematic of capillary gel-filling apparatus. Well degassed and mixed acrylamide solution was filled into the capillaries through the vacuum effect generated by a water aspirator.

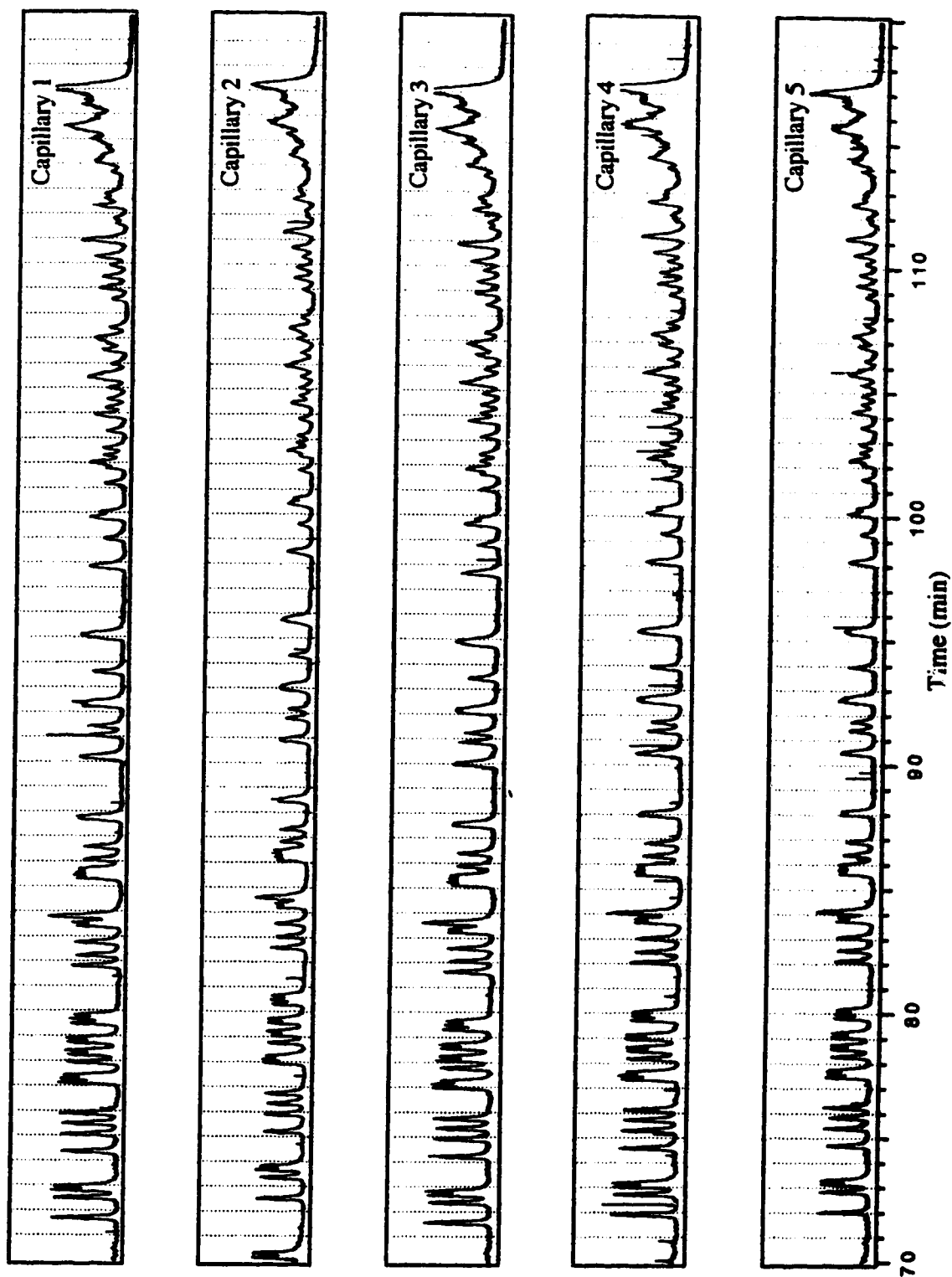
disconnected and the *in situ* polymerization was allowed to proceed overnight. These gel-filled capillaries were inspected under a microscope before installed into the sheath flow cuvette. These capillaries were trimmed 1- to 3-cm from the injection ends and then immersed in an electrode buffer ($1 \times$ TBE). Fig. 4. 12 presents the separation of C-terminated M13mp18 sequencing fragments in five capillaries. The separation was performed with 5%T, 0%C polyacrylamide, 7M urea, and $1 \times$ TBE buffer in 50- μ m ID, 150- μ m OD and 35 cm long capillaries at an electric field of 200 V/cm at room temperature. Migration times for the same fragment varied only by a few peak widths. Reproducible migration time results from careful attention to detail in preparation of the capillaries. It seems that the difference in migration time in free zone electrophoresis is largely originated from the surface effect, but this effect is effectively suppressed by the gel that covers the wall of the capillary.

Great success in separation of DNA sequencing fragments with this five capillary system has been achieved by using 5% non-cross-linked polyacrylamide as a separation medium. A series of 11 sample separations showed that between five capillaries, migration times differ only by 2 to 7 peak widths, corresponding to a 0.5 to 2% variation in migration time. The separation efficiencies typically range from 1.5 to 2.2 million theoretical plates for base 90).

However, the run-to-run variation in migration time can be 25% or even higher, depending on the initial condition of the gel. Usually, the electrophoretic current is a good indication of the gel condition. Fig. 4. 13 presents a relationship between the migration time as a function of electrophoretic current. Here, a total of 9 runs were made with five capillaries in each run. The eluting time is the average time when the reptation peak elutes from the capillary. The migration time is normalized to a capillary length of 35.0 cm. The current is the average current over the whole process of electrophoretic separation. While there are many conditions affecting the migration times, such as the actual gel concentration, the actual volume of the gel after degassing, the times each gel

Fig. 4.12 (pages 128-129). Electropherograms of the separation of FAM-labeled, C-terminated M13mp18 DNA sequencing fragments in five capillaries filled with 5%T linear polyacrylamide. Capillaries were 50 μm ID, 150 μm OD and 36.6 cm long. The sample was electrokinetically injected at 100 V/cm for 8 seconds and the electrophoresis was driven at an electric field strength of 200 V/cm.





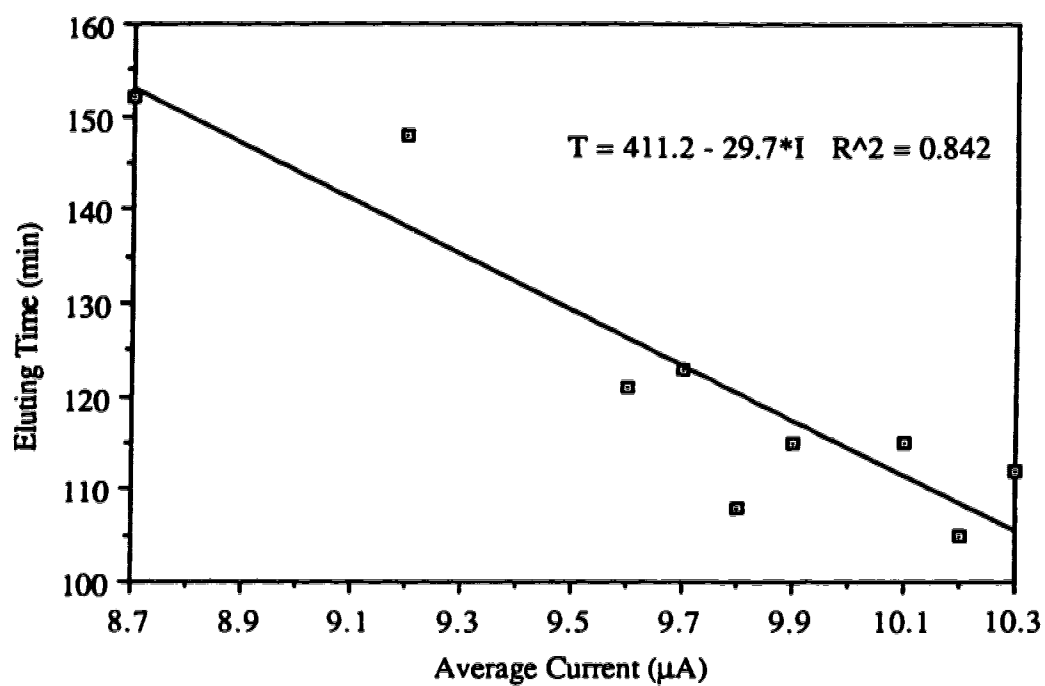


Fig. 4.13. The relationship between the eluting time of the reptation peaks and the average current of the electrophoresis. Capillaries were normalized to a length of 35.0 cm.

had been used, the buffer concentration and the room temperature, there is a tendency that low electrophoretic current would generate long migration time.

As has been presented in Chapter 2, the initial currents for newly prepared capillary gels are about the same (otherwise, the gels may not function properly in the later separation and are usually discarded). As the gel electrophoresis proceeds, an induced electric field is built up at the gel-buffer interfaces, which is opposite in direction to the applied electric field. This interfacial effect is caused by the buffer ions that possess different mobilities in gel and buffer phases. As the electrophoresis continues, a depletion zone is formed around the gel-buffer interfaces, much like a p-n junction in a semiconductor diode, and a field is built up across these interfaces to counteract the applied electric field. This argument has been proposed by Spencer in a series of studies [26-28] and by Swerdlow *et al.* along with some supporting phenomena [16]. Because of this induced electric field, the effective electric field, the electrophoretic current, and the migration speed of the DNA fragments will be reduced. To use each capillary gel more than once, time should be allowed for gel recovery. To speed up the recovery process, the polarity of the applied voltage can be reversed. To compare the sequencing speed, gels should be used that have the same conditions, particularly electrophoretic current.

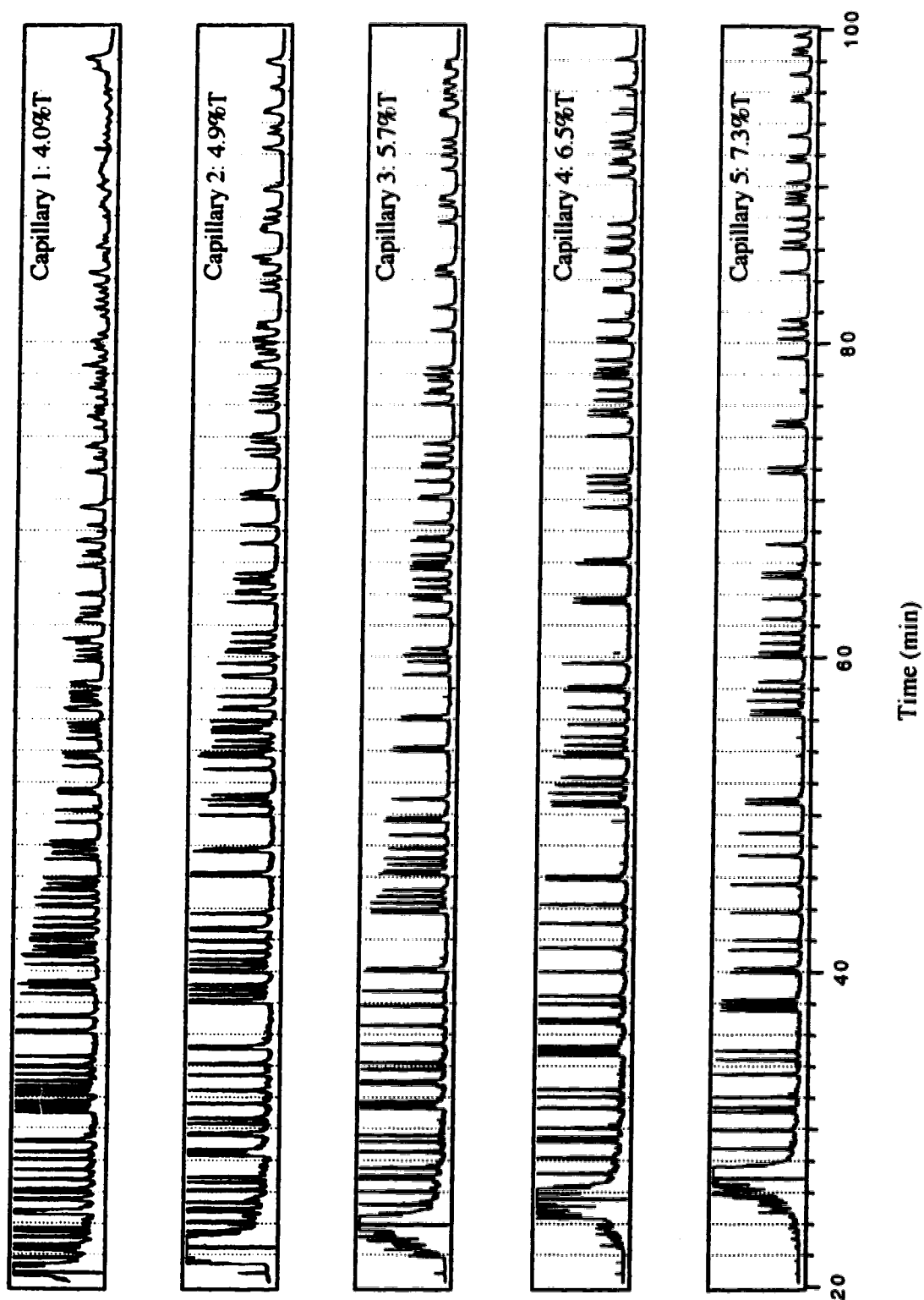
In the study of the relationship between the gel composition and separation efficiency, we have found that, at moderate electric field strength, longitudinal diffusion of the DNA fragments is the main source for band broadening [29]. Even though Equation 1-5 indicates that the separation efficiency is proportional to the applied voltage, fundamentally, the separation efficiency is related to the migration time, according to Equation 1-4. The longer the time a DNA fragment spend in the capillary, the more longitudinal diffusion occurs to lower the separation efficiency.

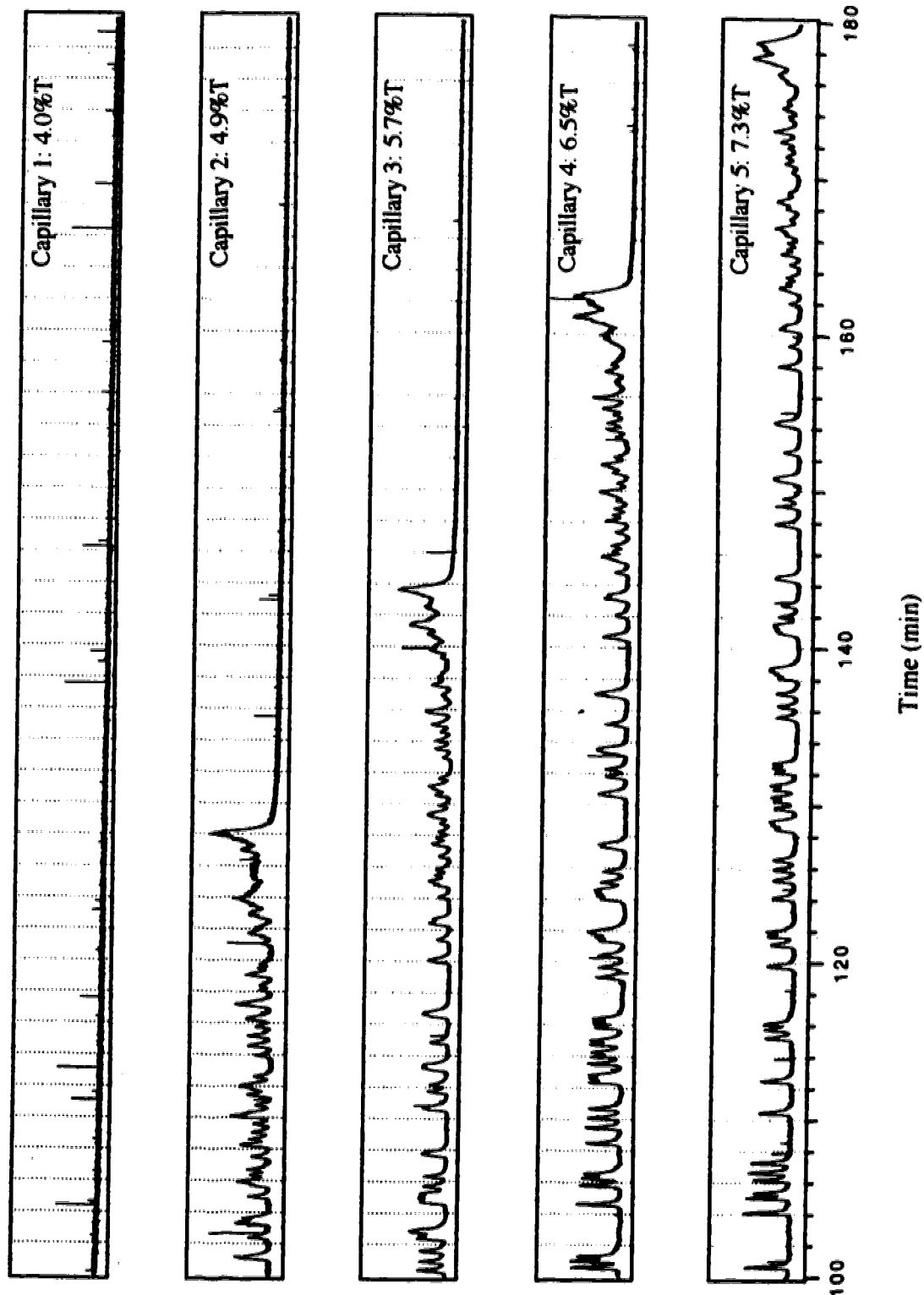
The similarity in migration time among the five capillaries is beneficial for parallel comparisons. For example, the effect of total acrylamide concentration in the absence of cross-linker was studied by using this five capillary instrument. In this study, a

4%T 0%C acrylamide was divided into five 5-ml aliquots and to each aliquot was added a certain amount of 40%T acrylamide stock solution to make a series of acrylamide concentrations: 4%T, 4.9%T, 5.7%T, 6.5%T, and 7.3%T. The polymer solutions were vacuumed into five capillaries simultaneously. These capillaries were then used to separate fluorescently labeled sequencing fragments at an electric field of 200 V/cm. Fig. 4. 14 presents the data. Increasing the gel concentration drastically slowed the sequencing rate; the biased reptation peak was observed at 100 minutes with 4%T polyacrylamide and at 185 minutes with 8%T polyacrylamide. As with cross-linked gels, the mobility of the longer fragments undergoing biased reptation increases linearly with total monomer concentration. Also note that the primer elution time increases slightly as the total monomer concentration increases. Last, low %T polyacrylamide physical gels generate more sequence information than the higher %T polyacrylamide.

Fig. 4. 15 presents the migration time versus DNA fragments generated from the observed data, Fig. 4.14. It is obvious that higher T% leads to longer separation time. Also we can note that these curves are more or less “S” shaped, particularly for higher %T. This means that the gels have limited resolving power for either very short or very long DNA fragments. Very short DNA fragments, which are smaller than the pore size of the gel structure, can move freely through the gels. With almost identical charge-to-size ratio, these small DNA fragments would have almost identical mobility. For very long fragments, however, the electrophoretic behavior is described by the biased reptation theory [30], or a modified Ogston sieving theory [31]. These theories predict that DNA fragments longer than a threshold will be stretched by the gel matrix and the applied electric field. As a result, the effective size of the fragment is decreased and its migration time is reduced. Fig. 4. 16, which is a relationship between the mobility and the fragment length, indicates that longer fragments tend to have very similar mobility and, eventually, the gels lose their power to separate them. Higher concentrations of the gel have smaller pore sizes, and the biased reptation occurs for smaller DNA fragments. From these data,

Fig. 4.14 (pages 134-135). Electropherograms of the separation of FAM-labeled, C-terminated M13mp18 DNA sequencing fragments in five capillaries filled with 4.0, 4.9, 5.7, 6.5, 7.3 %T linear polyacrylamide. Capillaries were 50 μm ID, 150 μm OD and 37.0 cm long. The sample was electrokinetically injected at 100 V/cm for 10 seconds and the electrophoresis was driven at an electric field strength of 200 V/cm.





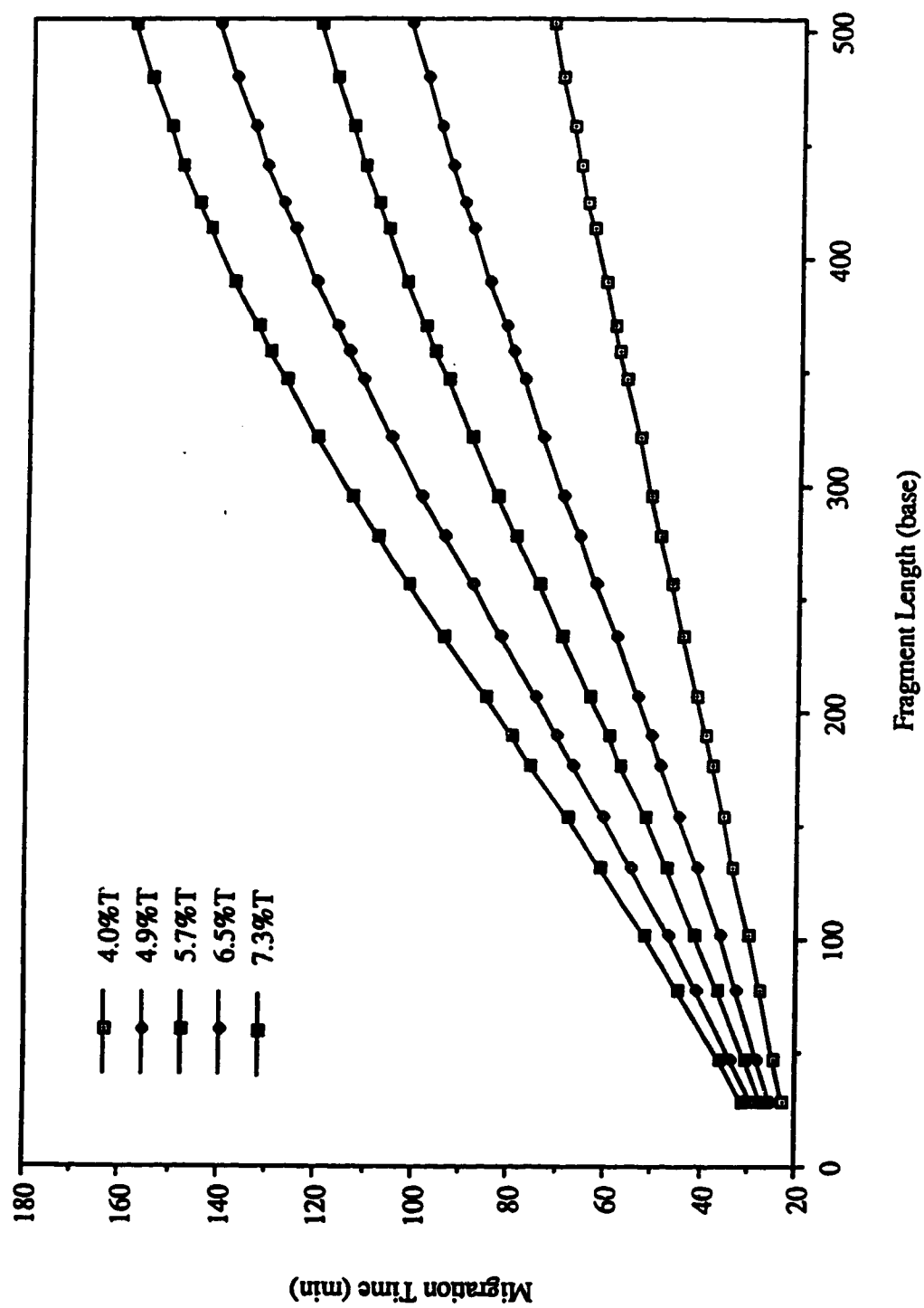


Fig. 4.15. Migration time as a function of DNA fragment length in different concentrations of linear polyacrylamide (%T). The polyacrylamide concentrations range from 4.0 to 7.3%.

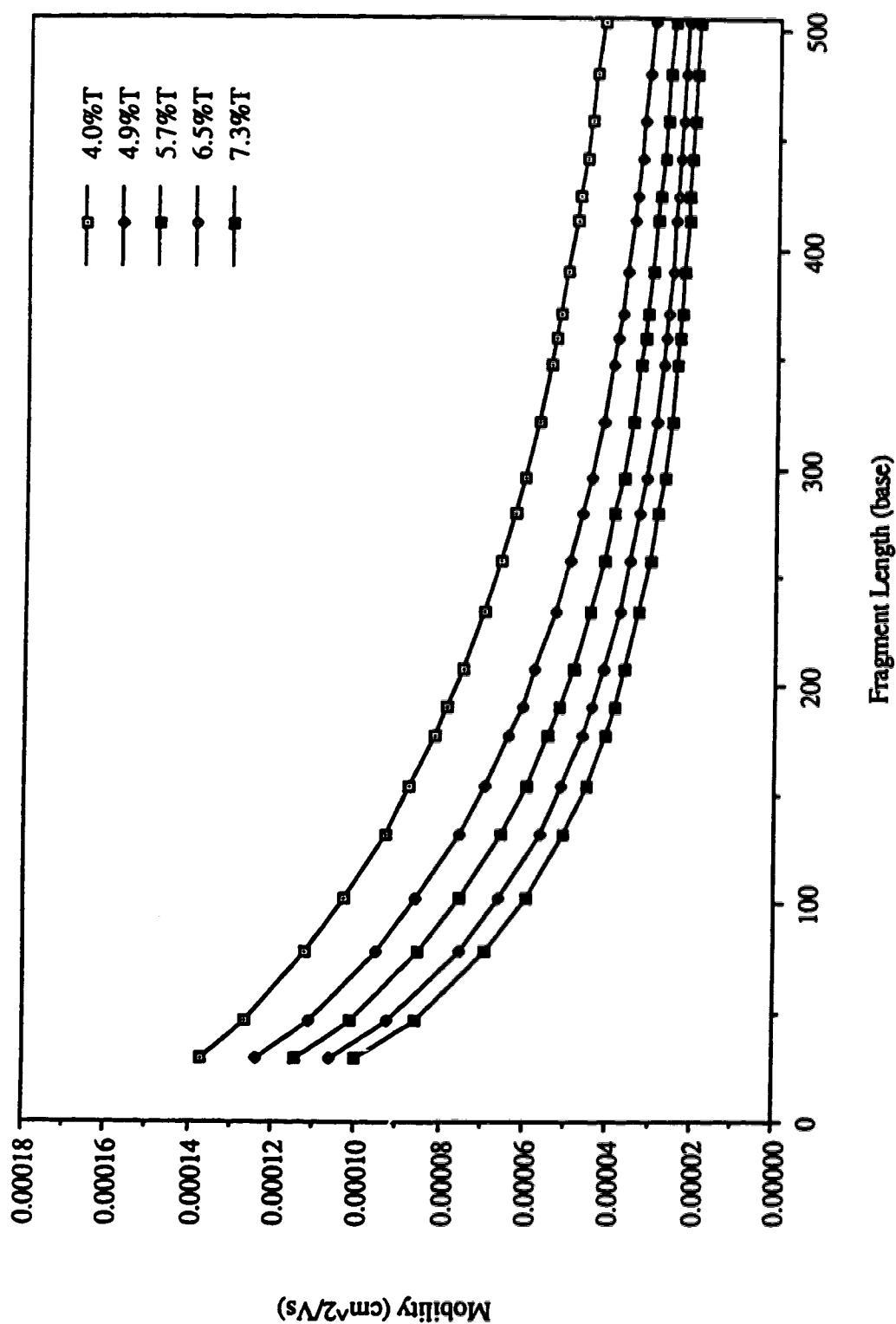


Fig. 4.16. Mobility of FAM-labeled DNA fragments in different concentrations of linear polyacrylamide (%T). The polyacrylamide concentrations range from 4.0 to 7.3%.

the mobility for a fragment of zero bases is estimated to be $1.6 \times 10^{-4} \text{ cm}^2/\text{V}\cdot\text{s}$, which is similar to $1.8 \times 10^{-4} \text{ cm}^2/\text{V}\cdot\text{s}$ that has been measured with conventional 5%C-cross-linked capillary gel [29].

As mentioned in Chapter 1, the gel behavior can be described by a Ferguson plot. Fig. 4. 17 shows this relationship for fragments with lengths: 29, 102, 207, 296, 390, and 503 bases. However, the linear relationship seems to only apply for smaller DNA fragments. The Ferguson plots become more curved for longer DNA sequencing fragments at higher %T. This result is in good agreement with the analysis performed by Tietz and Chrambach [32]. The curvature indicates that longer DNA fragments undergo biased reptation at higher %T. The curves should extrapolate to a common intercept, $\ln(\mu_0) = -8.34$, corresponding to a mobility in free solution of $2.2 \times 10^{-4} \text{ cm}^2/\text{V}\cdot\text{s}^{-1}$, which is equal to the value that has been obtained with conventional 5%C-cross-linked capillary gel [29].

For biased reptation, the relationship between mobility and the fragment length is given as follows

$$\mu = a \left(\frac{1}{N} + \frac{1}{N^*} \right) = \frac{a}{N} + \mu_{\infty} \quad (4-12)$$

where a is the slope of the line, N^* is the fragment length for the onset of biased reptation and μ_{∞} is the mobility of the longest fragments. Fig. 4. 18 presents the relationship between μ and $1/N$ for different gel concentrations. Biased reptation sets at ~ 250 bases for 4% gel while at ~ 130 bases for 7.3% gel.

It seems that 5%T linear polyacrylamide gel, used at room temperature, is adequate in terms of separation speed, separation efficiency and separation length. 4% linear polyacrylamide gel, which gives the highest separation speed and the longest sequence information, shows low separation efficiency for short DNA fragments. Improved separation efficiency for the short DNA fragments can be obtained by using

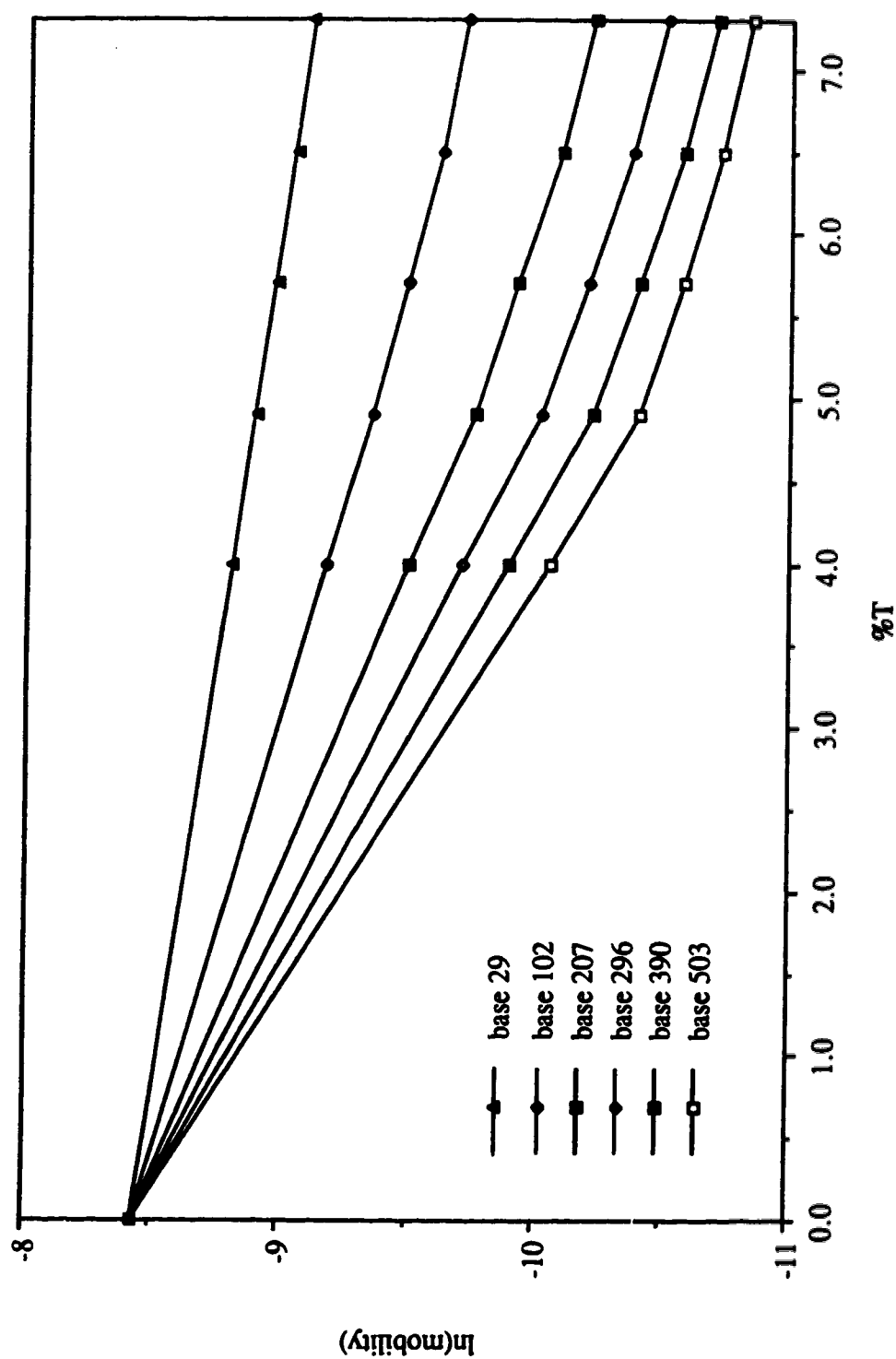


Fig. 4.17. Ferguson plots for DNA fragments. For base 29, there is a linear relationship: $\ln(\mu) = -8.43 - 0.095 \times (\%T)$ with $r = 0.999$. For other bases, the curves are extrapolated to $\ln(\mu_0) = -8.43$ at 0%T.

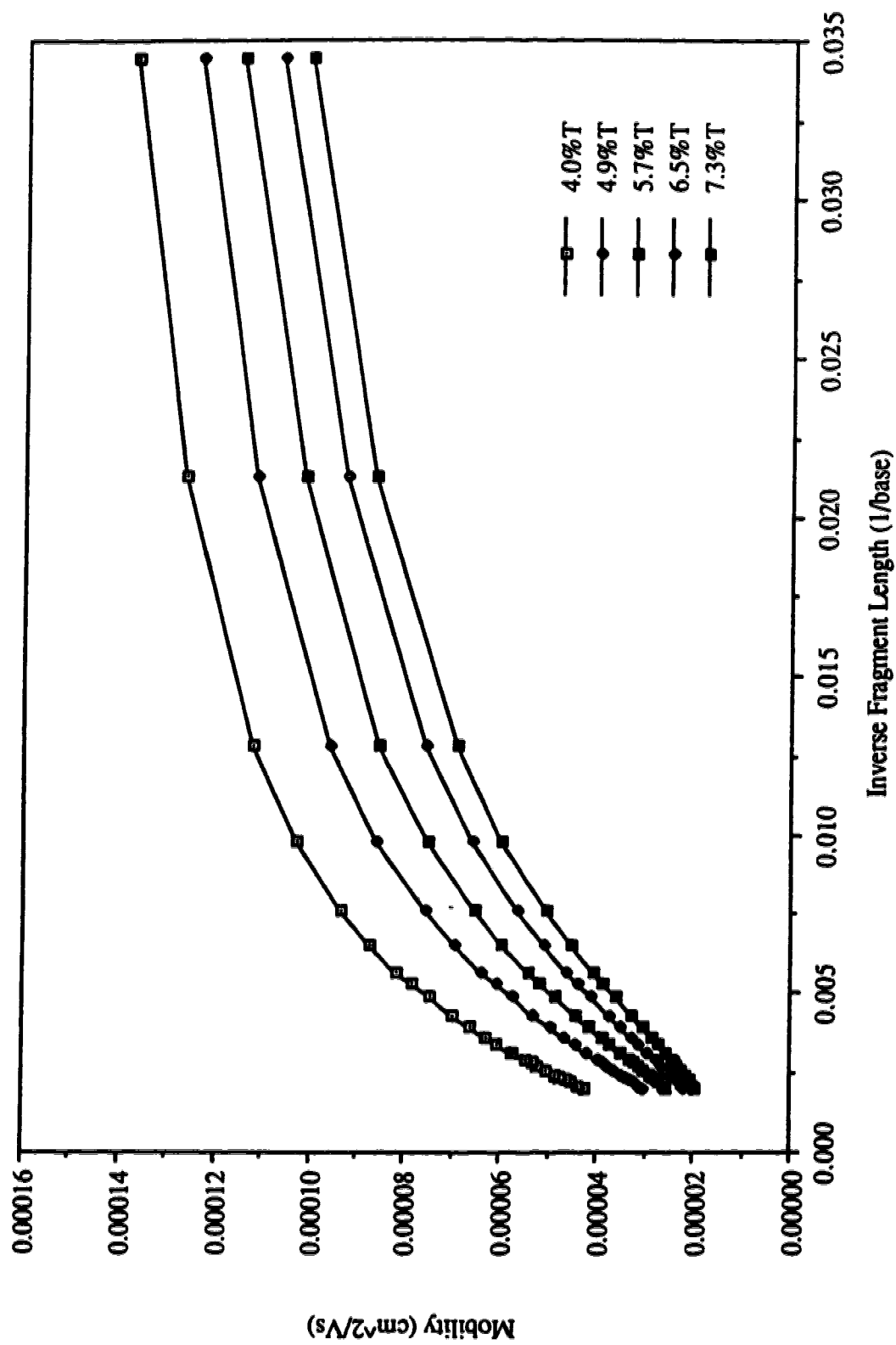


Fig. 4.18. Mobility of the DNA fragment as a function of inverse fragment length in linear polyacrylamide with different T%.

polyacrylamide gel with higher %T, however, the cost is longer separation time and shorter sequence information. The relatively shorter sequence information obtained by Huang *et al.* [13] and by Pentoney *et al.* [22] may be due to the use of even higher %T (9%T for the former and 10%T for the latter) linear polyacrylamide for the separation.

4.4.3. DNA SEQUENCING

The PRISM chain terminators developed by ABI are labeled with FAM, ZOE, LOU, and NAN (see section 1.3.3. in Chapter 1), and produce good signals and separation when used with T7 DNA polymerase [5]. Because the fluorescence dyes have relatively widely spaced absorbance and emission spectra, two lines from an argon ion laser (488 nm) and a green helium-neon laser (543.5 nm) were used to excite fluorescence and four interference filters were used to isolate emission at 540, 560, 580 and 610 nm, corresponding to the wavelengths longer than the maximum emission of the dyes. Excitation at 488 nm was for FAM and ZOE and excitation at 543.5 nm was for LOU and NAN. Although the laser power is lower at 543.5 nm (2.0 mW) than at 488 nm (4.0 mW), the intensity generated at the beam waist may be sufficient to photobleach many fluorophores [33,34,1].

A filter wheel was rotated to transmit sequentially fluorescence intensity in four spectral bands to the photodetector array. A 150-mer peak passed detector region in 6 s; the filter wheel was rotated at 1.5 Hz, providing 9 data points per spectral channel per peak which was sufficient to reconstruct the sequencing data. A sector wheel was rotated at the same rate as the filter wheel to block alternately the two laser beams; excitation at 488 nm was associated with measurement of emission at 540 and 560 nm while excitation at 514.5 nm was associated with measurement of emission at 580 and 610 nm. This combination of excitation/emission channels not only eliminated interference by Rayleigh scatter at 543.5-nm excitation wavelength with the 540-nm emission channel

but also interference by Raman scatter due to 488-nm excitation wavelength with the 580-nm emission channel. Synchronization of the sector and filter wheels was achieved through use of stepper motors and a single digital controller (Fig. 4.19).

Data acquisition was not synchronized with the rotation of the filter set. According to the speed of the rotation wheel, the data acquisition rate should be $4 \times 1.5 = 6$ points per second per channel. The actual data acquisition rate, however, was ~ 24 points per second per channel, 4 times higher than the rotation speed of the wheels. In other words, each filter has included 4 data points when passing through the photodetectors.

If $Y_i(p)$ represents the raw data in the form of a one dimensional array, where i represents the separation channel ($i = 0, 1, 2, 3$ and 4), p represents the data point number in the array. Two steps were involved in data processing. First, the maxima, $W_i(p)$, were selected:

$$\text{If } Y_i(p-1) < Y_i(p) \geq Y_i(p+1)$$

$$\text{Then } W_i(p) = Y_i(p)$$

Each selected data is only a 30-ms counts. Fig. 4.20a shows the raw data and selected maxima taken from 52.85 to 53.20 minutes. Second, the selected data were divided into four groups, Ch_{i0} , Ch_{i1} , Ch_{i2} and Ch_{i3} , each group representing fluorescence signal detected in one spectral channel

$$Ch_{i0}(p) = W_i(4p)$$

$$Ch_{i1}(p) = W_i(4p+1)$$

$$Ch_{i2}(p) = W_i(4p+2)$$

$$Ch_{i3}(p) = W_i(4p+3)$$

Fig. 4.20b presents a section of the data processed in this way, showing the four spectral information for base identification. The data is taken from base 214 to 221. The orange, red, green and blue traces correspond to fluorescence signals centered at 540, 560, 580, and 610 nm, respectively. Bases T, C, A and G can be identified from the spectral

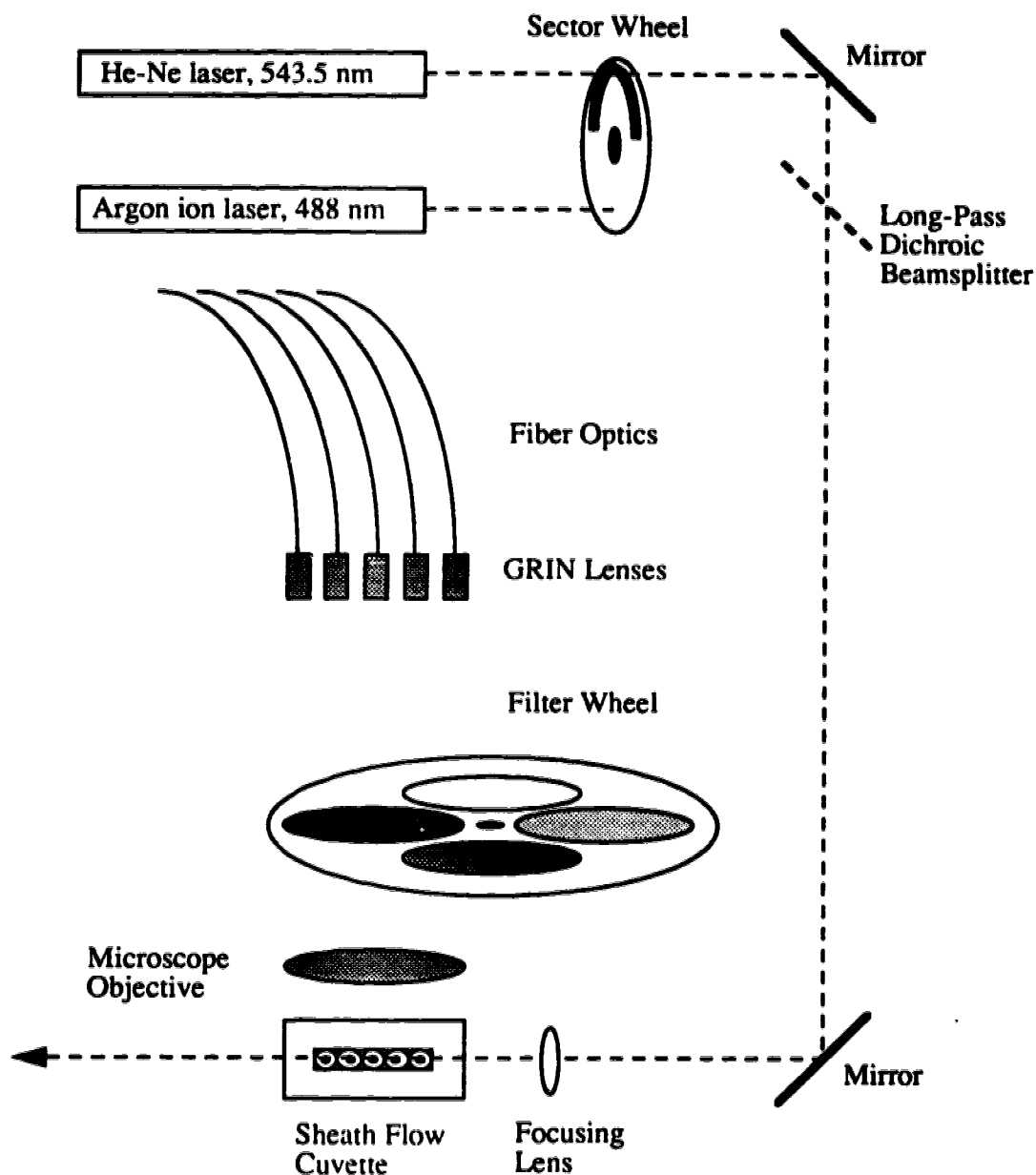


Fig. 4.19. Laser-induced fluorescence detector for four spectral-channel sequencing. The sector wheel alternately transmits the two laser lines and the filter wheel synchronize the transmission of fluorescence at wavelength centered at 540, 560, 580 and 610 nm. See text for details.

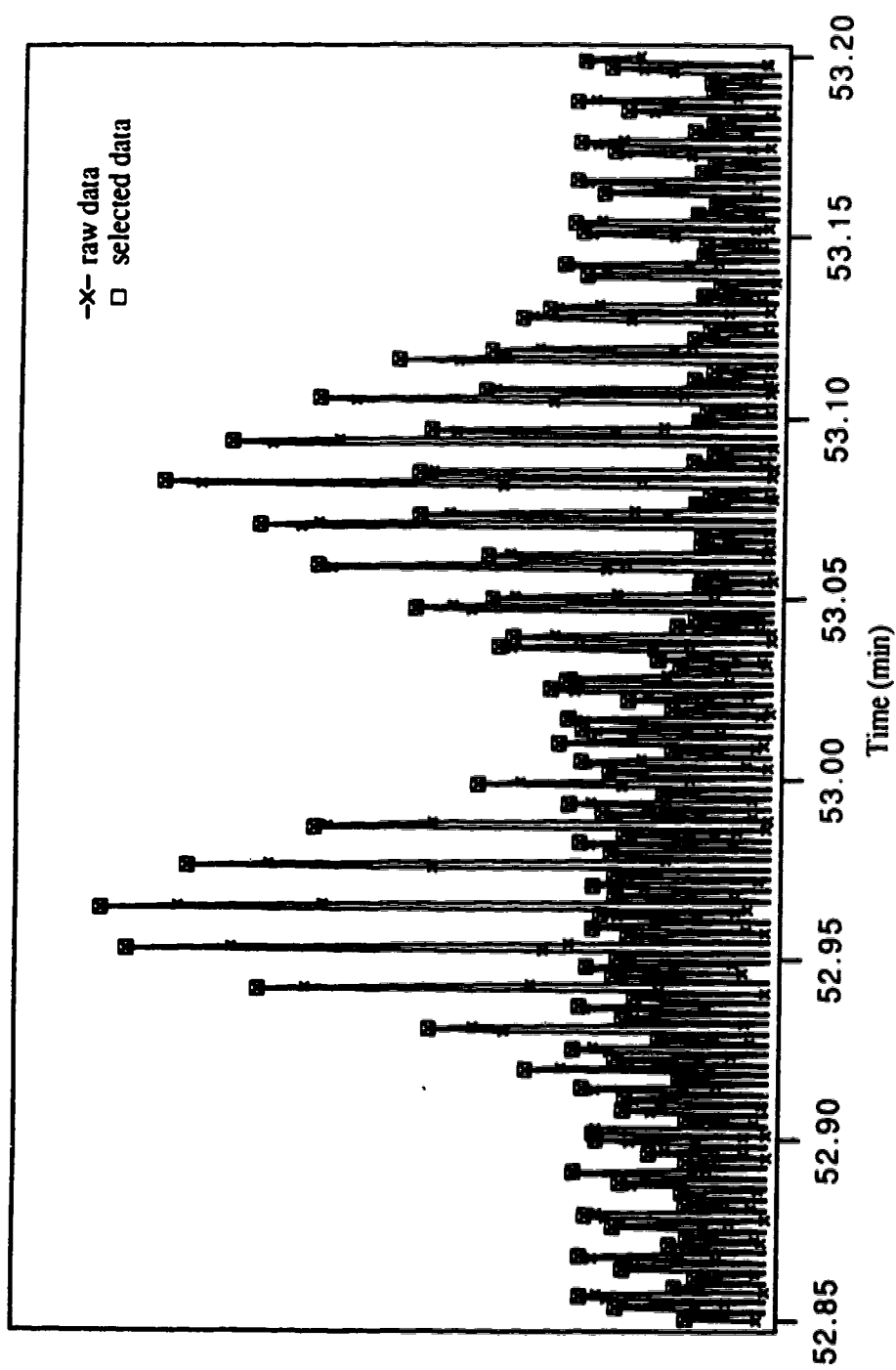
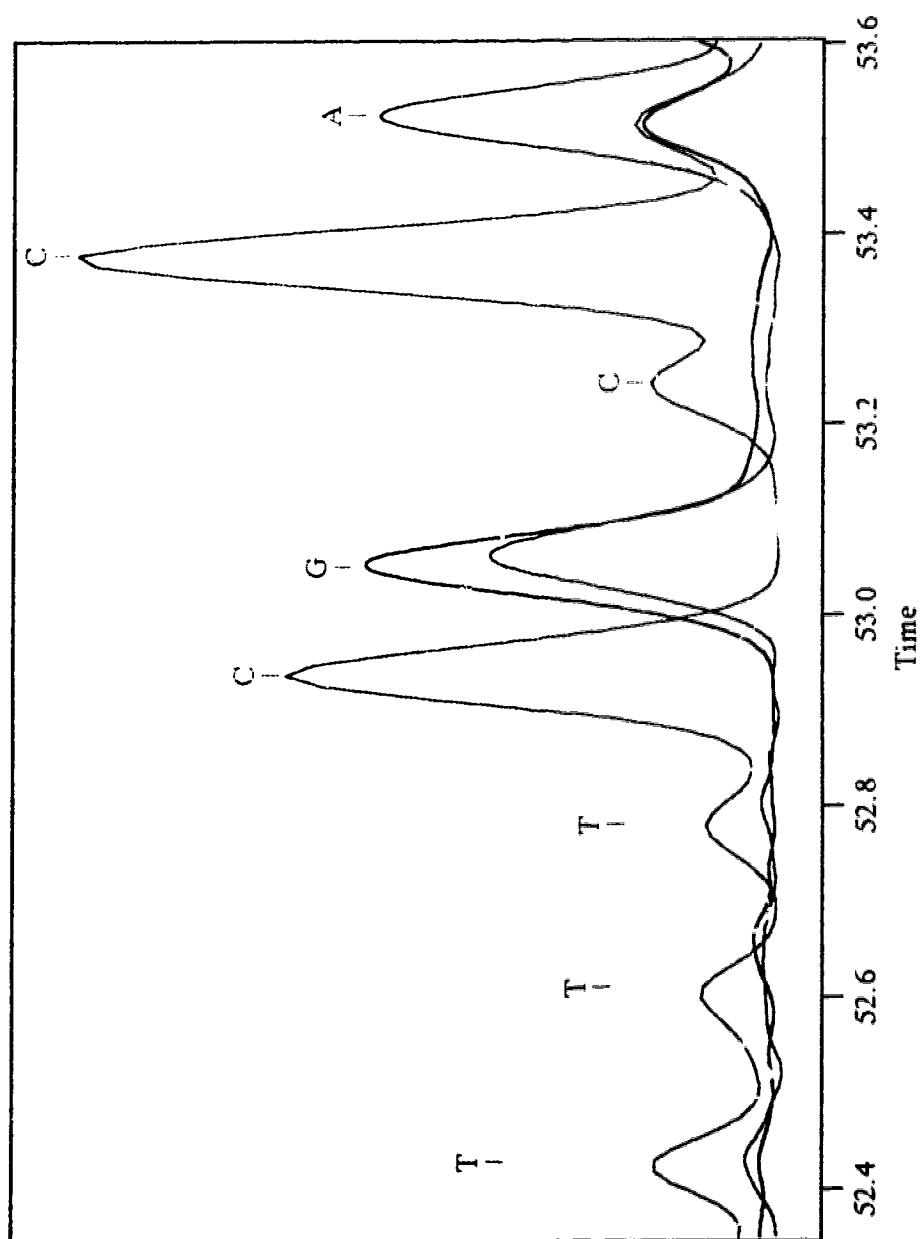


Fig. 4.20. (a) Sequencing raw data and the selected maxima taken from 52.85 to 53.20 minutes from one of the five separation channels.

Fig. 4.20. (b) A section of the processed data showing the four spectral information, Ch_{i0} , Ch_{i1} , Ch_{i2} and Ch_{i3} , for base identification. The data is taken from base 214 to 221. The orange, red, green and blue traces correspond to fluorescence signals centered at 540, 560, 580, and 610 nm, respectively. For each peak, the highest trace in orange indicates a T, the highest trace in red indicates a C, the highest trace in green indicates an A and the highest trace in blue indicates a G.



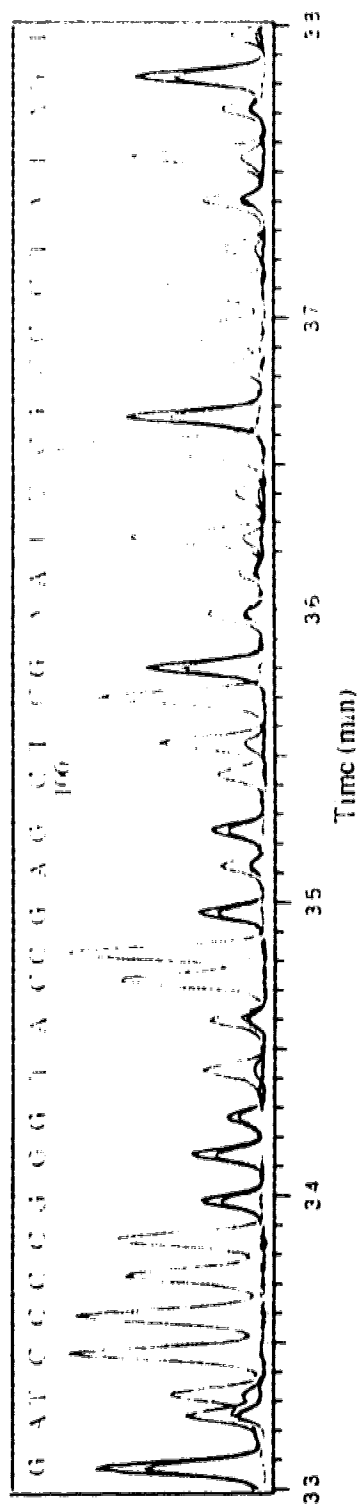
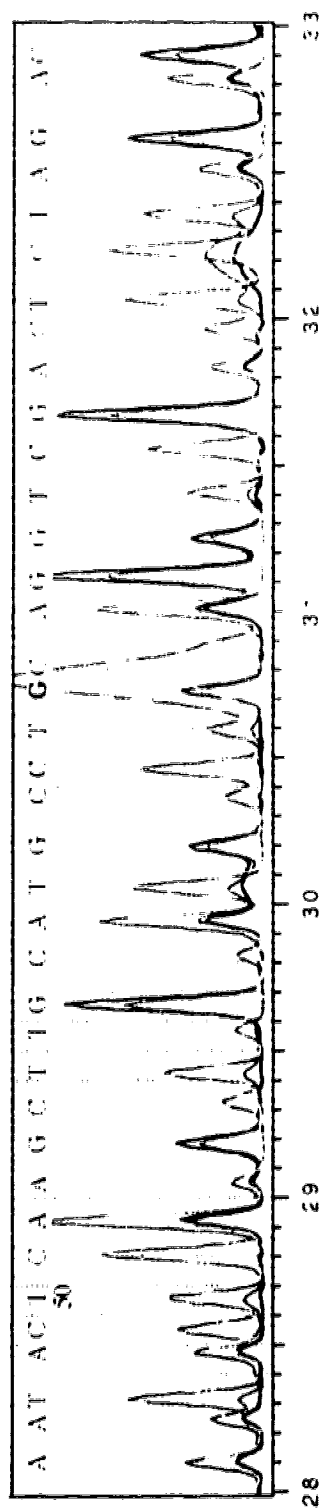
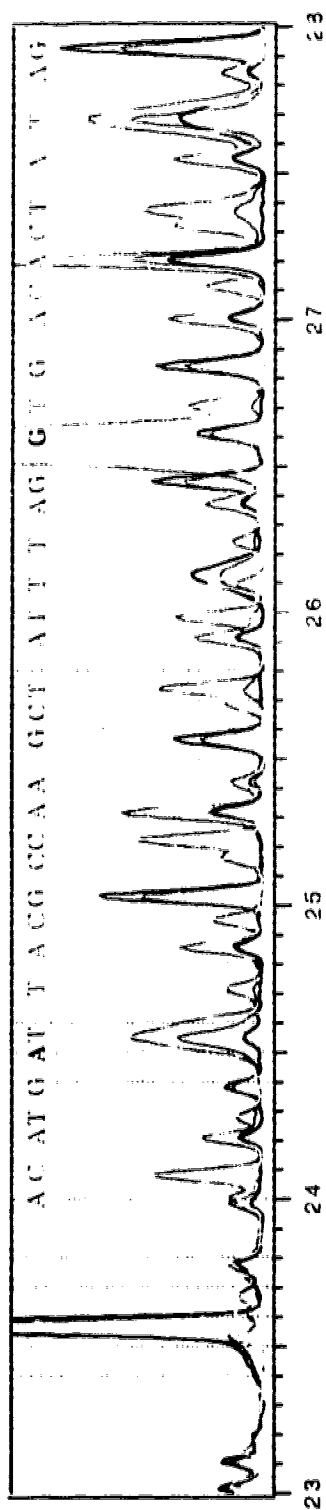
information. For bases A and G, the fluorescence signals are multiplied by a factor of 2 in order to compensate for the lower excitation power from the He-Ne laser and to make the peak height more uniform for the four bases.

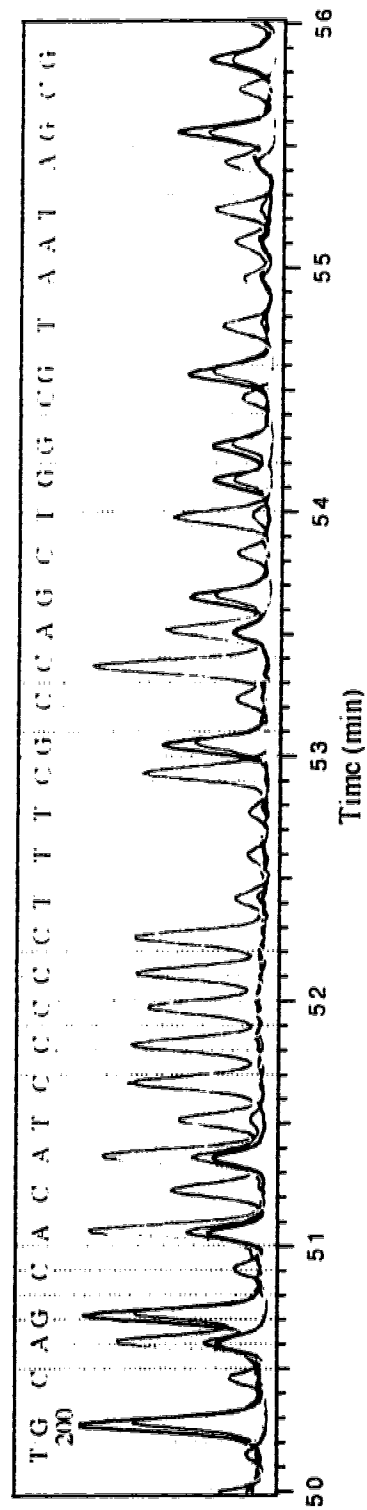
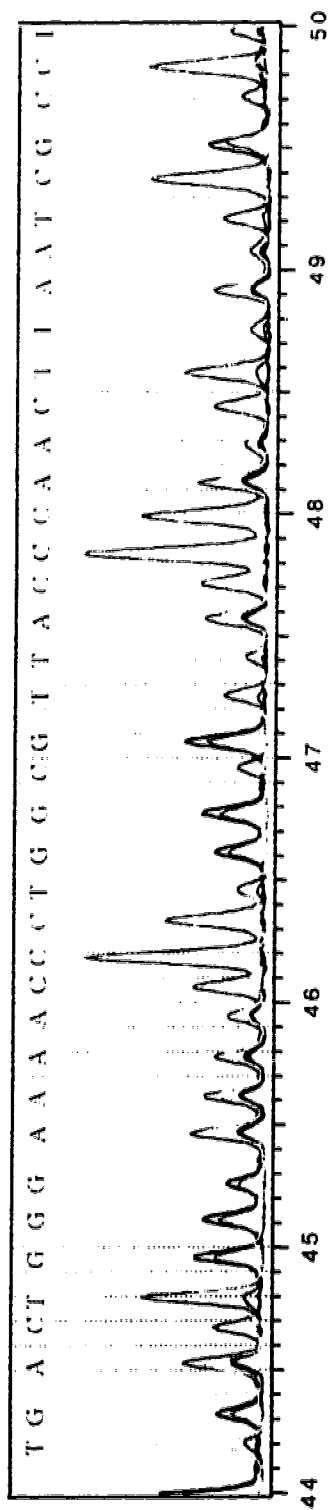
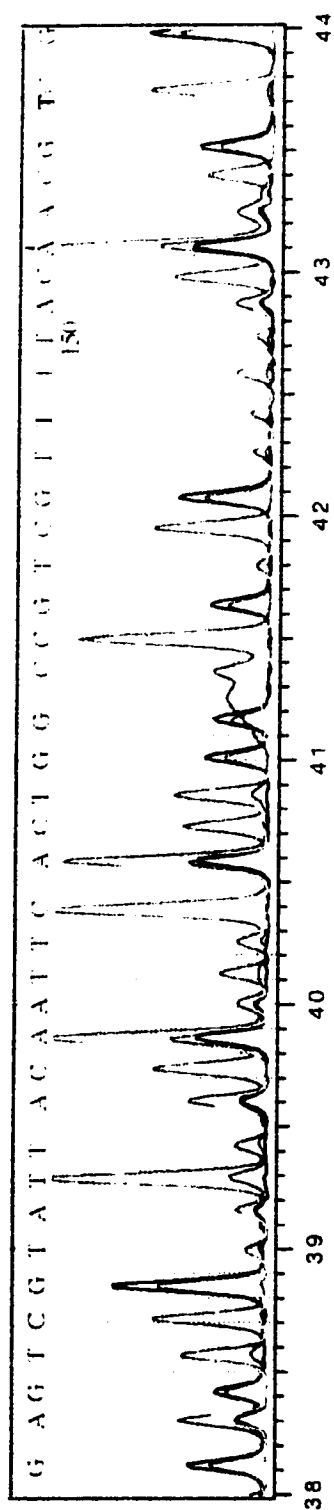
It should be noted that instead of the filter set required by ABI's instrument for the dye-terminators (530, 545, 560 and 580 nm [5]), very good spectral resolution has been found by using the alternative filter set (540, 560, 580 and 610 nm), which is the combination used for ABI's dye-primers with the dyes being FAM, JOE, TAMRA and ROX. This difference may be due to the fact that different photodetectors have been used in the two instruments. In this five capillary instrument, SiAPD is the photodetector, which has a higher quantum efficiency at longer wavelengths, whereas in ABI's instrument, PMT is the photodetector, which has lower quantum efficiency at longer wavelength (see Table 3.1 for comparison).

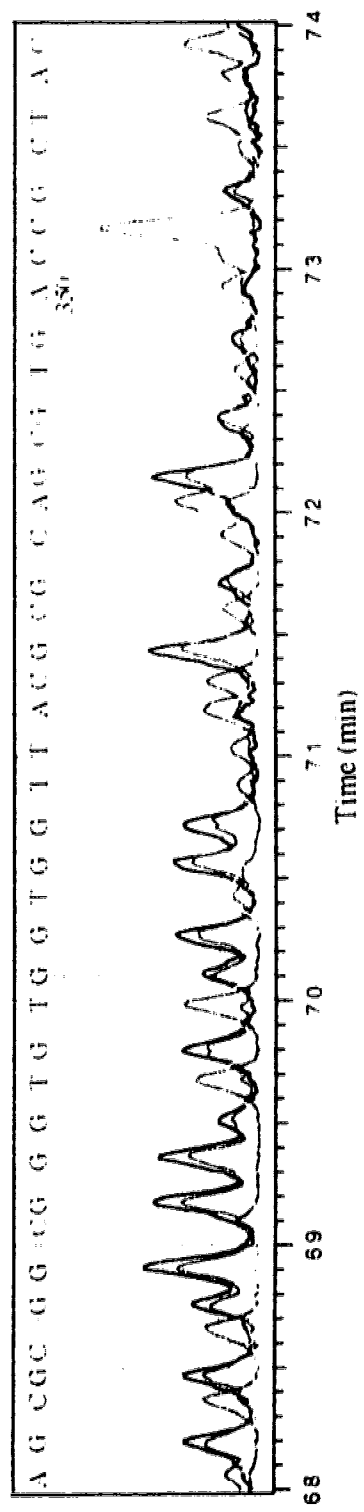
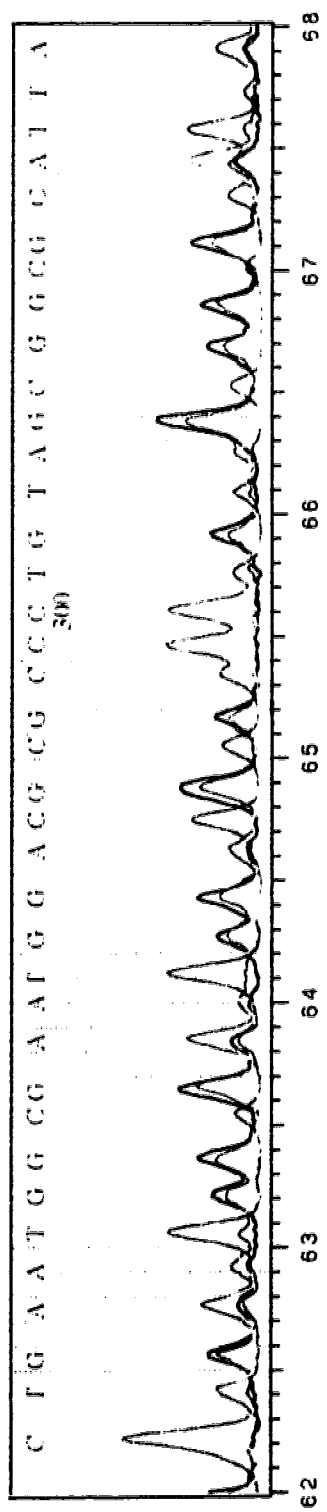
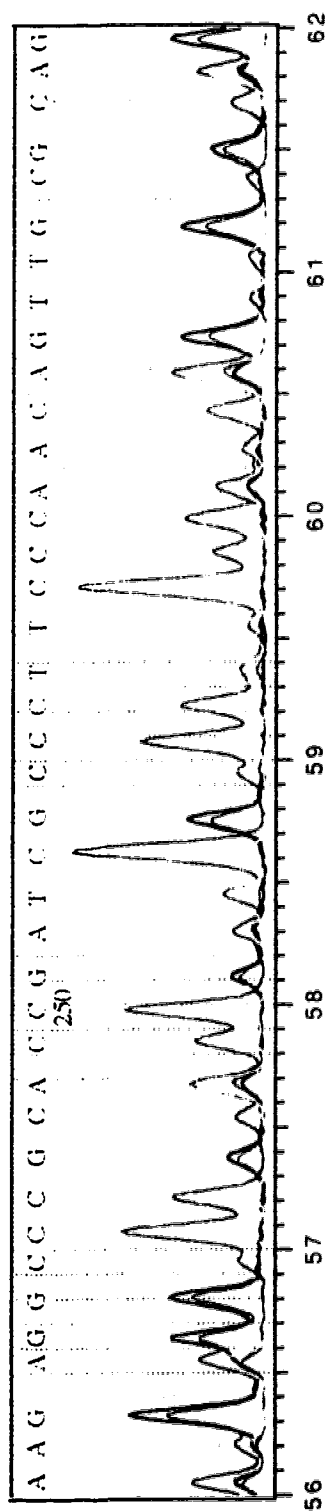
Fig. 4.21 presents the sequence generated from the single-stranded PGEM-REV template from one of the five separation and detection channels. The separation of the sequencing sample was at 200 V/cm in 5%T linear polyacrylamide gel, producing a sequencing rate of 400 bases per hour per capillary, or 2000 base per hour per instrument. In each capillary, over 500 bases of sequence was determined from the data; comparison allowed peaks to be identified beyond this point. From bases 50 to 220 the peaks were nearly baseline resolved. Beyond base 220, the separation began to suffer from the effect of biased reptation and the resolution began to degrade. The small broad red trace between bases 140 and 141 was probably due to the sample contamination. A few minor compressions can be found at bases 14 and 15, bases 32 and 33, bases 66 and 67 and at bases 156 and 157. The rest of the sequence can be read correctly according to the spectral information shown in Fig. 4.20a, corresponding to an accuracy of ~99% within the range from 10 to 500 bases.

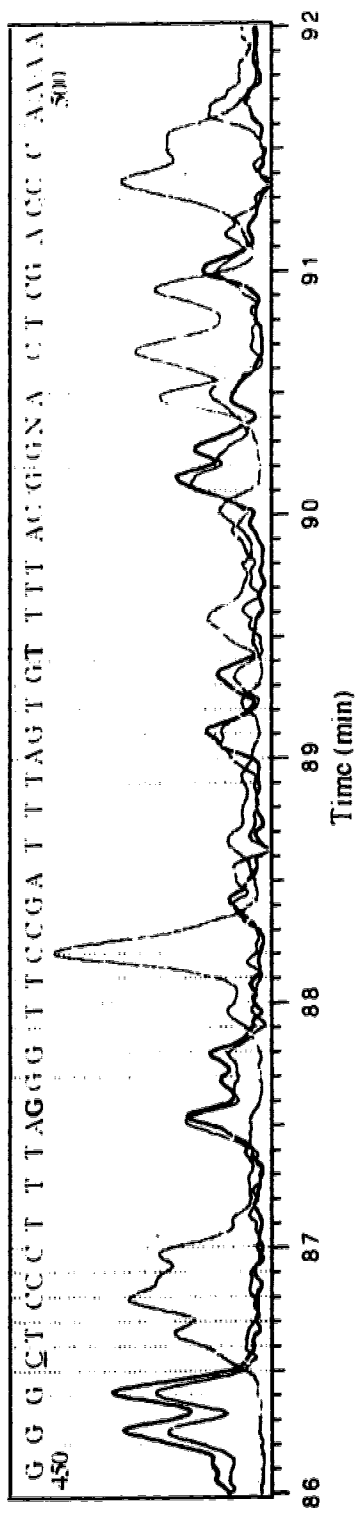
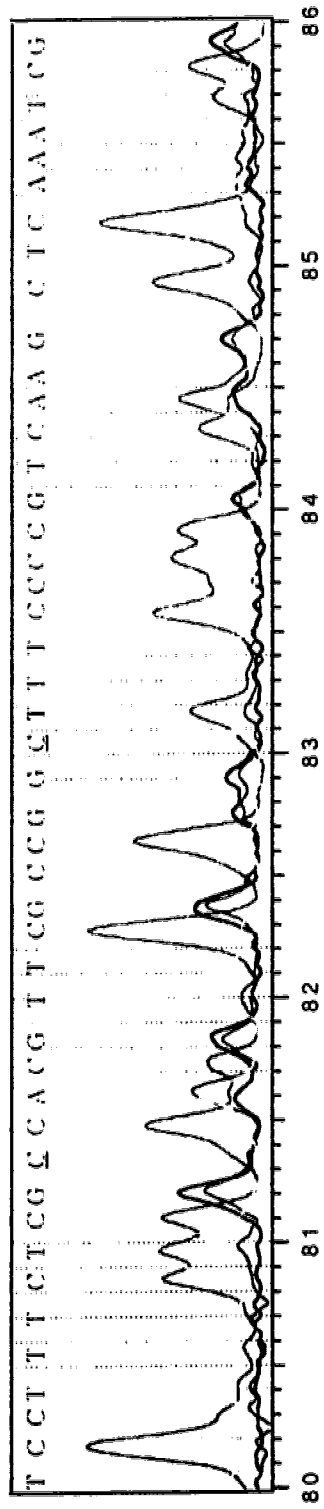
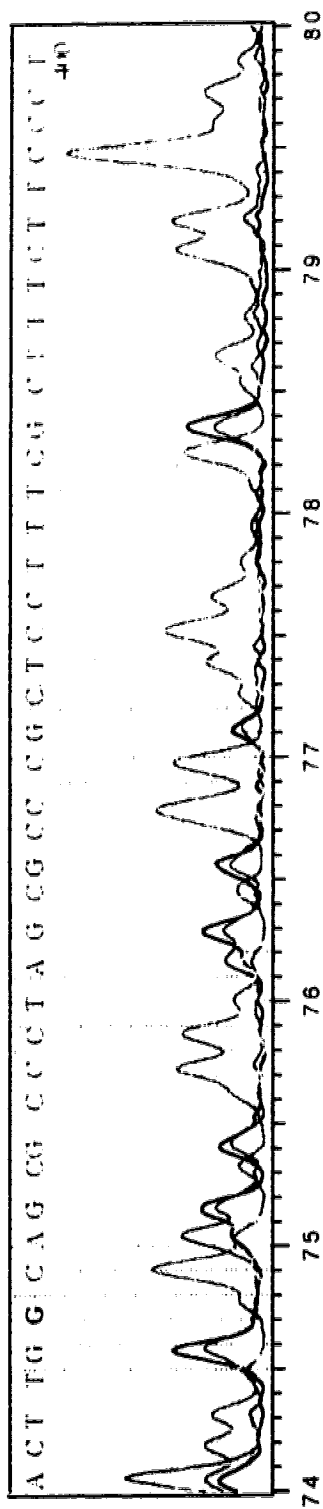
Comparison of this result with that reported by Huang *et al.* [23], which is the only published sequencing data from multiple capillary instrument, demonstrates that

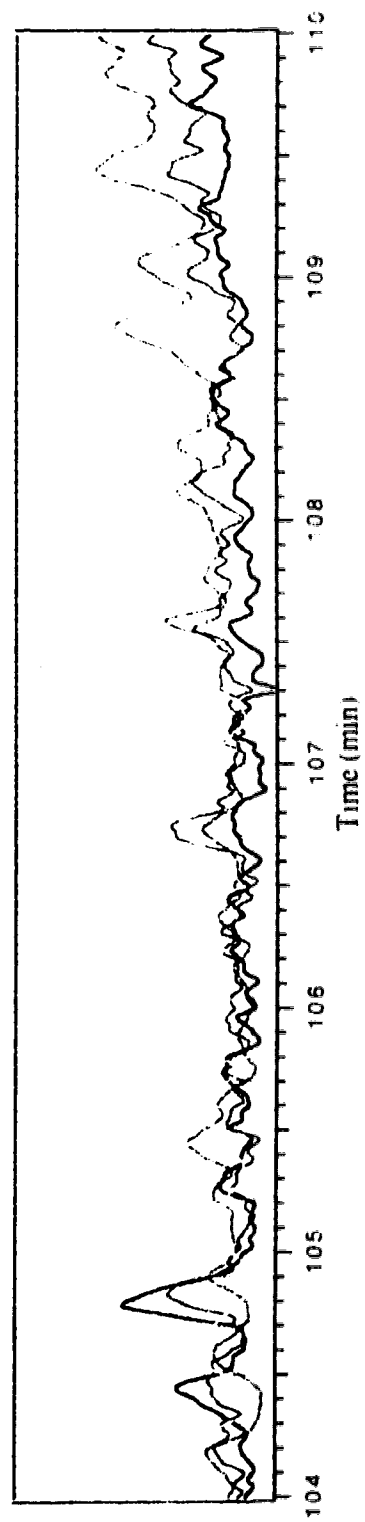
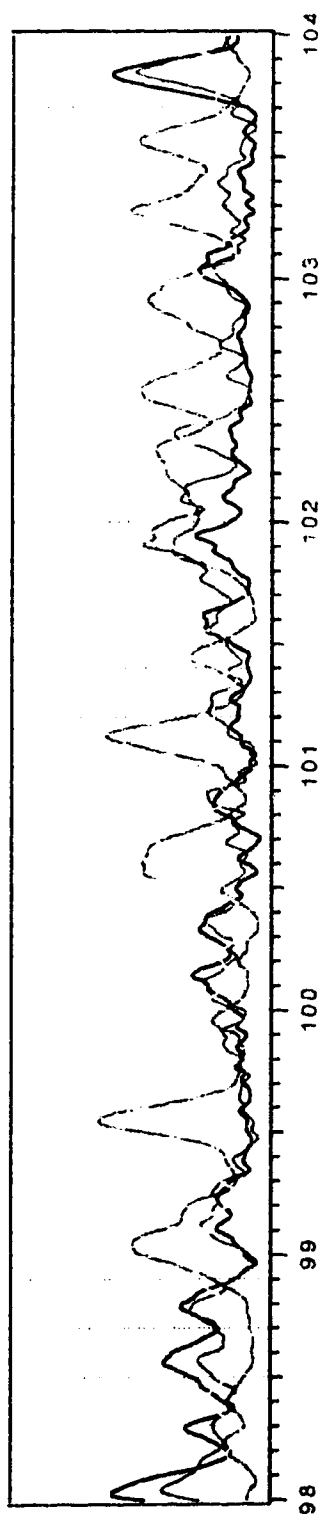
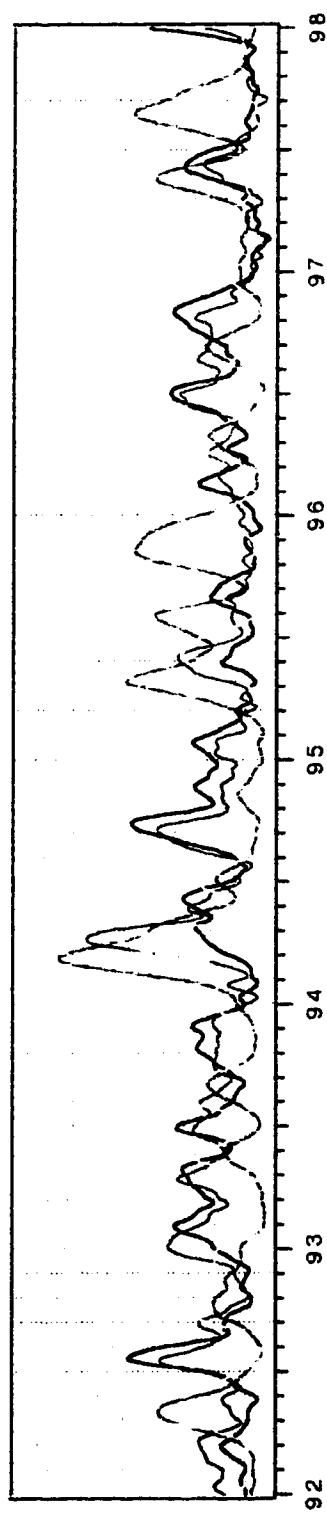
Fig. 4.21 (pages 149-153). Four spectral, dye-terminator DNA sequencing. The sample was generated from a control template, single-stranded PGEM-REV, in one representative capillary. Time is indicated from 23 to 110 minutes and base numbers are indicated at 50-base interval to the corresponding peaks. The sequence is printed above. Bold letters indicate base calling errors. Underlined letters are the callings which were given as N's by ABI.











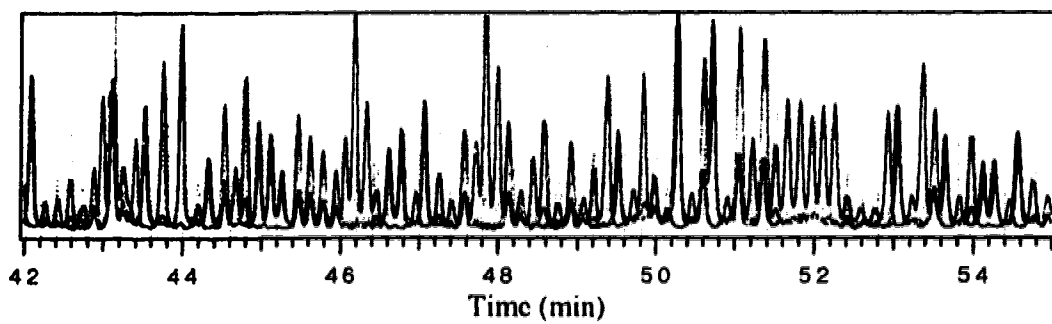
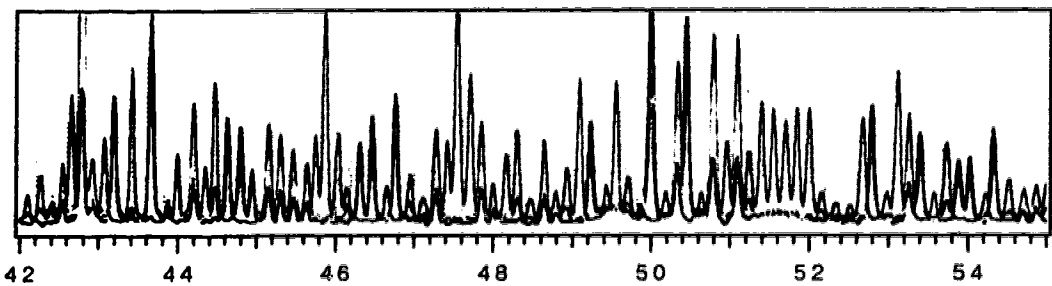
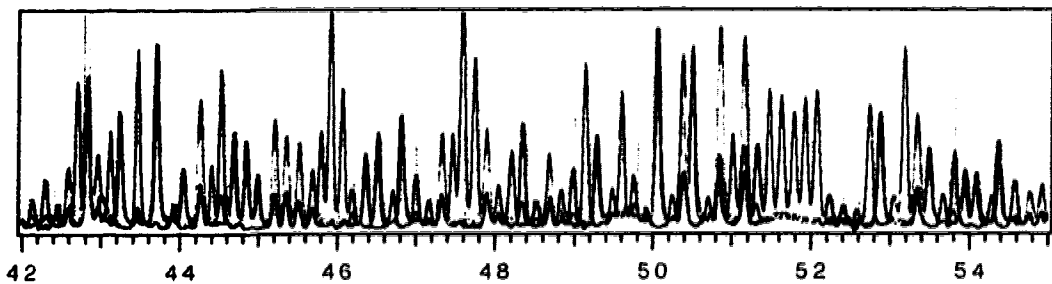
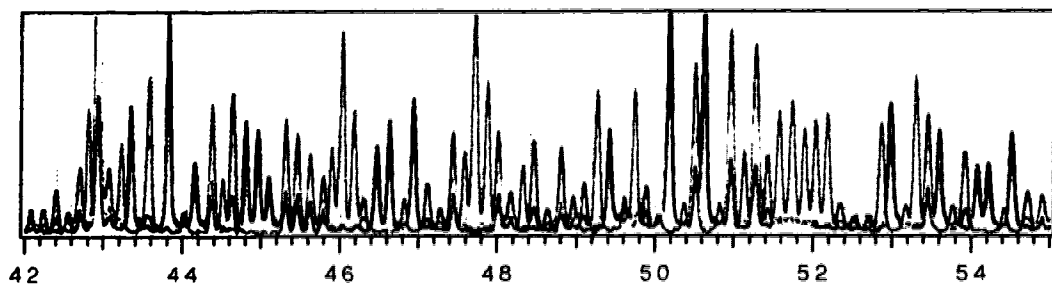
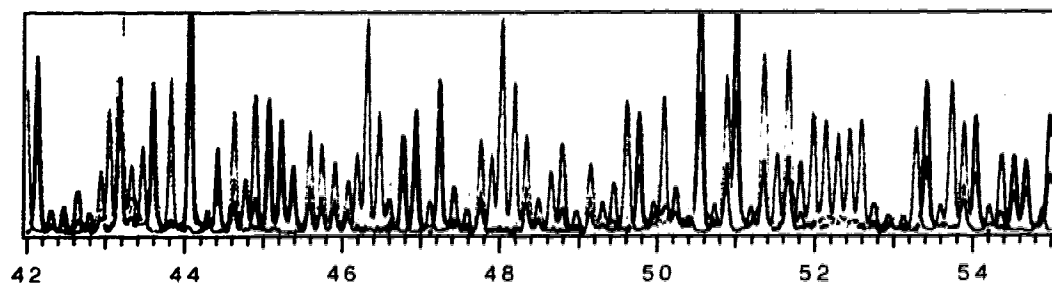
high sensitivity of the instrument is essential in long sequence determination. In the report from Huang, on-column fluorescence detection was achieved by scanning the capillary array. Capillaries were 100 μm ID. However, the sequence determination was limited to 350 bases in length; beyond this point, the fluorescence signals vanished in the system noise.

Fig. 4.22 shows a section, from 42 to 55 minutes, of the separation of the single-stranded PGEM-REV sequencing sample obtained from the five separation and detection channels. Very similar separation patterns were obtained from the five channels. The migration time for the same fragment varied at most by four peak widths.

4. 5. CONCLUSIONS

In DNA sequencing, it is beneficial, in terms of time and cost, to have long sequence determination in a single run. Longer fragments are lower in signal level and tend to co-migrate. High sensitivity of the instrument and the low-concentration of linear polyacrylamide gel are valuable in achieving desirable separation for DNA sequencing in multiple capillary format. We can anticipate that parallel operation of multiple capillaries will be a strong candidate for large-scale genome sequencing.

Fig. 4.22. A section of simultaneous separation in five capillaries of DNA sequencing fragments generated from a control template—single-stranded PGEM-REV. Time scale is indicated at the bottom.



REFERENCES

1. Wu, S. and Dovichi, N.J., *J. Chromatography*, 1989. **480** : p. 141-155.
2. Kambara, H. and Takahashi, S., *Nature*, 1993. **361** : p. 565-566.
3. Cannon, B., Gardner, T.S., and Cohen, D.K., *Applied Optics*, 1986. **25** (17): p. 2981-2983.
4. Cohen, D.K., Little, B., and Luecke, F.S., *Applied Optics*, 1984. **23** (4): p. 637-640.
5. Lee, L.G., Connell, C.R., Woo, S.L., D., C.R., McArdle, B.F., Fuller, C.W., Halloran, N.D., and Wilson, R.K., *Nucleic Acids Research*, 1992. **20** (10): p. 2471-2483.
6. Wu, S. and Dovichi, N.J., *Talanta*, 1992. **39** (2): p. 173-178.
7. Li, L.-Q. and Davis, L.M., *Rev. Sci. Instrum.*, 1993. **64** (6): p. 1534-1529.
8. Rocheleau, M.-J. and Dovichi, N.J., *J. Microcol. Sep.*, 1990. **4**: p. 449-453.
9. Lux, J.A., Yin, H.-F., and Schomburg, G., *J. High Resolution Chromatography*, 1990. **13** : p. 436-437.
10. Dolnik, V., Cobb, K.A., and Novotny, M., *J. Microcol. Sep.*, 1991. **3**: p. 155-159.
11. Cohen, A.S., Najarian, D.R., Paulus, A., Guttman, A., Smith, J.A., and Karger, B.L., *Proc. Natl. Acad. Sci. USA*, 1988. **85** : p. 9660-9663.
12. Baba, Y., Matsuura, T., Wakamoto, K., Morita, Y., Nishitsu, Y., and Tsuhako, M., *Anal. Chem.*, 1992. **64** : p. 1221-1225.
13. Huang, X.C., Quesada, M.A., and Mathies, R.A., *Anal. Chem.*, 1992. **64** : p. 967-972.
14. Chen, D.Y., Swerdlow, H.P., Harke, H.R., Zhang, J.Z., and Dovichi, N.J., *J. Chromatography*, 1991. **559**: p. 237-246.
15. Sweedler, J.V., Shear, J.B., Fishman, H.A., Zare, R.N., and Scheller, R.H., *Anal. Chem.*, 1991. **63** : p. 496-502.

16. Swerdlow, H., Dew-Jager, K.E., Brady, K., Grey, R., Dovichi, N.J., and Gesteland, R., *Electrophoresis*, 1992. **13** : p. 475-483.
17. Ruiz-Martinez, M.C., Berka, J., Belenkii, A., Foret, F., Miller, A.W., and Karger, B.L., *Ana. Chem.*, 1993. **65** : p. 2851-2858.
18. Tietz, D., Gottlieb, M.H., Fawcett, J.S., and Chrambach, A., *Electrophoresis*, 1986. **7**: p. 217-220.
19. Pulyaeva, H., Wheeler, D., Garner, M.M., and Chrambach, A., *Electrophoresis*, 1992. **13** : p. 608-614.
20. Tietz, D., Aldroubi, A., Pulyaeva, M., Gusszczyński, T., Garner, M.M., and Chrambach, A., *Electrophoresis*, 1992. **13** : p. 614-616.
21. Sudor, J., Foret, F., and Bocek, P., *Electrophoresis*, 1991. **12** : p. 1056-1058.
22. Pentoney, S.L.J., Konrad, K.D., and Kaye, W., *Electrophoresis*, 1992. **13**: p. 467-474.
23. Huang, X.C., Quesada, M.A., and Mathies, R.A., *Anal. Chem.*, 1992. **64** : p. 2149-2154.
24. Chiari, M., Nesi, M., Fazio, M., and Righetti, P.G., *Electrophoresis*, 1992. **13** : p. 690-697.
25. Guttman, A., Wanders, B., and Cooke, N., *Anal. Chem.*, 1992. **64** : p. 2348-2351.
26. Spencer, M., *Electrophoresis*, 1983. **4**: p. 36-41.
27. Spencer, M., *Electrophoresis*, 1983. **4**: p. 41-45.
28. Spencer, M. and Kirk, J.M., *Electrophoresis*, 1983. **4**: p. 46-52.
29. Harke, H.R., Sue, B., Zhang, J.Z., Rocheleau, M.J., and Dovichi, N.J., *J. Chromatography*, 1992. **608**: p. 143-150
30. Lumpkin, O.J., Déjardin, P., and Zimm, B.H., *Biopolymers*, 1985. **24** : p. 2181.
31. Luckey, J.A. and Smith, L.M., *Electrophoresis*, 1993. **14** : p. 492-501.
32. Tietz, D. and Chrambach, A., *Electrophoresis*, 1992. **13**: p. 286-294.
33. Hirshfeld, T., *Applied Optics* , 1976. **15** : p. 3135-3139.

34. Mathies, R.A. and Peck, K., *Anal. Chem.*, 1990. **62** : p. 1786-1791.

Chapter 5

Conclusions and Future Work

5.1. CONCLUSIONS

Fluorescence detection based on sheath flow cuvette has been successfully applied to single and one-dimensional array capillary systems. Fluorescein mass detection limits lower than 2 zeptomoles were demonstrated for both single and multiple capillary systems. PMT and SiAPD both are sensitive photon detectors with SiAPD being more sensitive than PMT.

Crosslinked polyacrylamide gels have been successfully applied for DNA fragment separation in single capillary system. High electric field strength could generate high sequencing speed. The electric field strength can not, however, be increased without bound. At room temperature, perhaps 500 V/cm is the ultimate limit for capillary electrophoresis with crosslinked polyacrylamide gel as the separation medium. Care must be taken in manipulating the applied electric field and capillary gel in order to achieve high performance in DNA sequencing.

It is the non-crosslinked polyacrylamide polymer, at least for time being, that makes the multiple capillary system applicable for convenient DNA fragment separation. The flexibility of this polymer network allows a care-free performance in DNA sequencing at moderate electric field strength. Uniform gel composition can be achieved by simultaneously filling acrylamide solution into an array of capillaries with a home-made apparatus.

This multiple capillary system, with its increased speed and decreased sample volume and concentration, parallel comparison capability, and versatility, could find

many applications in DNA sequencing, PCR product analysis, oligonucleotide purity assessment, clinic diagnosis and forensic analysis, to mention just a few. A patent has recently been applied for this system.

So far two major problems for DNA sequencing on a multiple capillary system have been solved: we have built an instrument that generates superior detection limits in fluorescence detection and we have demonstrated that low-concentration, non-cross-linked polyacrylamide is valuable in DNA fragment separation. However, the future still seems very challenging.

5.2. FUTURE WORK

To enhance the performance of the multiple capillary system, ways must be found that synchronize the data acquisition speed with that of the rotating filter wheel in DNA sequencing of the four-dye samples. One way to do this is to operate SiAPD in the analog mode instead of the Geiger mode. The data acquisition must be triggered by an external signal. Operating SiAPD in analog mode instead of Geiger mode may generate another advantage—expanded linear range, even though the internal amplification will be lower.

Developing dye-labeled samples with longer excitation and emission wavelengths may increase the detection sensitivity because the background will be lower and the sensitivity of the SiAPD will be higher at longer wavelength.

Making the gel replaceable seems very important [1,2] for at least three reasons. First, it will reduce the cost for DNA sequencing because each set of capillaries can be used many times. Second, fresh gels can be used for every run, eliminating possible gel contamination. Third, with gel being replaced each time, frequent alignment of the instrument is no longer necessary.

In this thesis, the DNA sequence was determined manually by judging the peak shape and relative fluorescence intensity from the individual spectral channels. However,

it is necessary that software be developed to automatically determine the DNA sequence from generated data.

Capillaries may be arranged in a two-dimensional array in a modified sheath flow cuvette. In this way, hundreds of gel-filled capillaries can be accommodated in a single sequencing machine for the use where massive sequence analysis is required, such as the Human Genome Project. Fig. 5.1 shows a schematic of the sheath flow cuvette fluorescence detector for two-dimensional capillary array and four-color fluorescence detection.

In the fluorescence detector, capillaries terminate in a sheath flow cuvette. The sheath stream draws the sample from each capillary, generating a two-dimensional array of isolated sample streams in the cuvette. Fluorescence is generated from all sample streams with an elliptically shaped laser beam, which is aligned down-stream from the ends of the capillaries. The sheath flow cuvette is tilted at 45° so that fluorescent spots do not overlap when imaged onto the CCD.

A cylindrical lens telescope, which is routinely used in flow cytometry instruments to produce slit-scans of single cells [3], is used to generate the elliptical shaped laser beam. A moderate power argon ion laser, operated at 488 and 514.5 nm, is used as the light source. Two filter wheels are used to choose the excitation wavelength as well as emission passband. Two line filters are used to alternately transmit 488 and 514.5 nm laser lines. A set of four bandpass filters is used to sequentially transmit the fluorescence from the sample streams, just like the one used in chapter 4. The rotation of the two filter wheels are synchronized so that fluorescence excited with the 488 nm laser line is detected with the two short-wavelength filters and fluorescence excited with the 514.5 nm line is detected with the two long-wavelength filters.

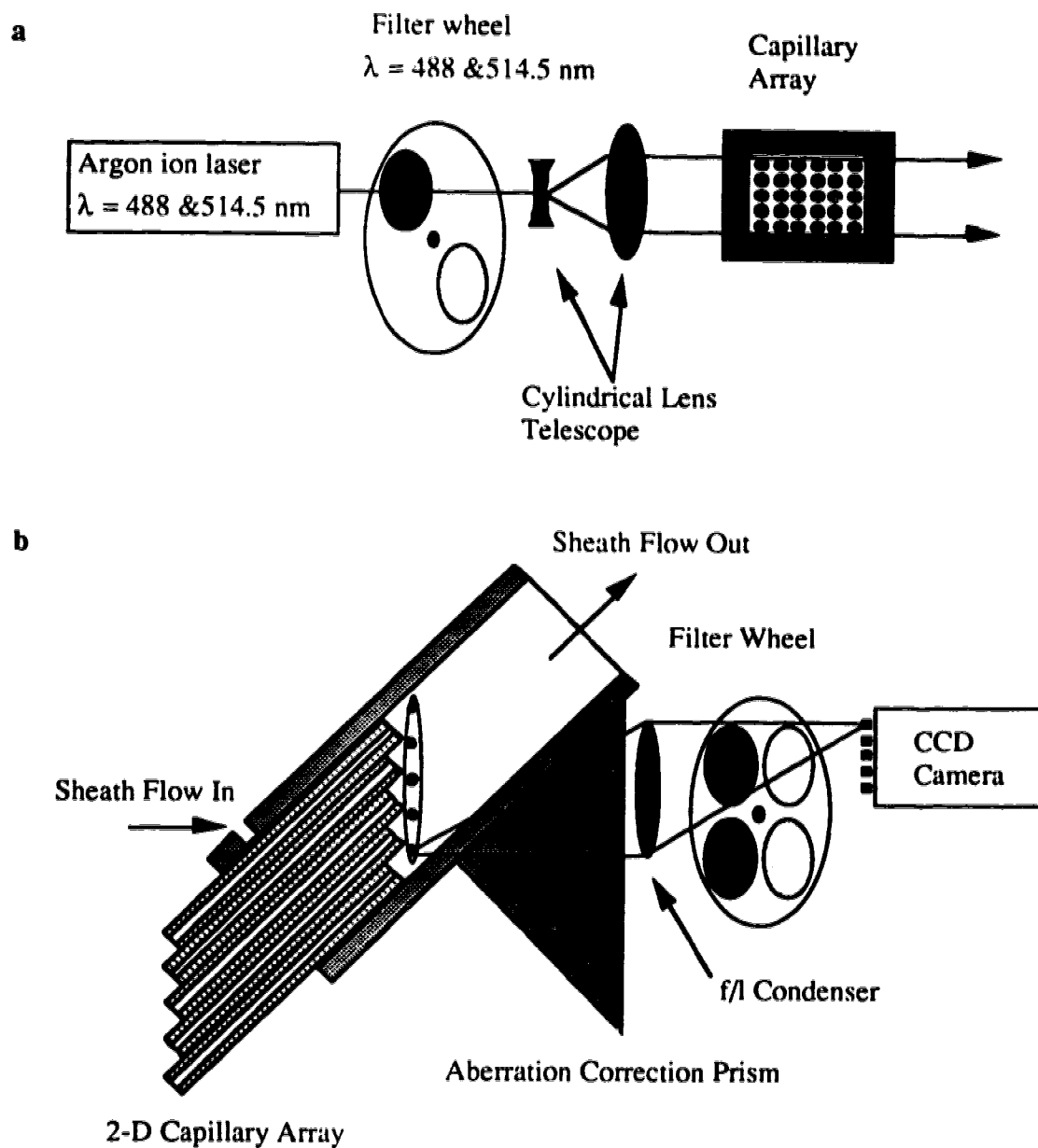


Fig. 5.1. A schematic diagram of a sheath flow cuvette fluorescence detector for two-dimensional capillary array and four-spectral fluorescence detection. (a) Excitation. (b) Detection.

REFERENCES

1. Ruiz-Martinez, M.C., Berka, J., Belenkii, A., Foret, F., Miller, A.W., and Karger, B.L., *Ana. Chem.*, 1993. **65** : p. 2851-2858.
2. Werner, W.E., Demorest, D.M., and Wiktorowicz, J.E., *Electrophoresis*, 1993. **14** : p. 759-763.
3. Johnston, R.G., *Rev. Sci. Instrum.*, 1985. **55** : p. 691.